

Em 28 de março de 2000, um grupo de pesquisadores paulistas publicou um artigo na revista científica norte-americana *Proceedings of the National Academy of Sciences (PNAS)* informando a seus pares do exterior que haviam desenvolvido uma metodologia capaz de identificar fragmentos de genes expressos (ou ativos). Era uma forma alternativa e complementar à técnica convencional de obtenção de ESTs – pedaços de genes ativos que, em inglês, são denominados *expressed sequence tags* –, que tinha sido criada em 1991 nos Estados Unidos.

Concebida por dois cientistas que então trabalhavam na filial paulista do Instituto Ludwig de Pesquisa sobre o Câncer, Andrew Simpson e Emmanuel Dias Neto, a metodologia brasileira recebeu o nome de Orestes (*Open Reading Expressed Sequence Tags*), que, em português, significa algo como etiquetas da fase aberta de leitura de seqüências expressas. Desde então, por desvendar a região central dos genes, o Orestes tornou-se uma ferramenta útil, sobretudo na busca por genes que se mostram ativos apenas em poucos tipos de tecidos, em projetos brasileiros e internacionais que estudam o genoma de organismos ou a expressão de genes envolvidos em diversas enfermidades. Foi usado também no estudo do genoma do verme *Schistosoma mansoni*, causador da esquistossomose, recentemente concluído.

Na edição impressa de 11 de novembro passado, o mesmo *PNAS* estampou em suas páginas um artigo que faz um balanço do uso conjunto do Orestes e da técnica mais tradicional de geração de ESTs no estudo de genes ligados a uma das mais desafiadoras doenças humanas, o câncer. O

trabalho foi escrito por cerca de 140 pesquisadores do Brasil, dos Estados Unidos, da Europa e da África do Sul que participaram de duas grandes iniciativas que analisam seqüências expressas em tumores: o Programa Genoma Humano do Câncer, financiado pela FAPESP e pelo Instituto Ludwig, e o *Cancer Genome Anatomy Project (CGAP)*, bancado pelo Instituto Nacional do Câncer dos Estados Unidos.

No artigo de seis páginas, a equipe multinacional de cientistas faz um apanhado dos resultados alcançados pelos projetos brasileiro e norte-americano, que, a partir de amostras de células sadias e com tumores, produziram informações detalhadas sobre o conjunto de genes que são ativados em tecidos extraídos de sete partes do corpo humano: pulmão, mama, cérebro, cabeça e pescoço, cólon, útero e rim. Com as amostras de células de cada uma dessas regiões, foram geradas pelo menos 100 mil seqüências expressas. “Mostramos que, nesse conjunto de tecidos, há uma variabilidade surpreendente no uso de genes, o que levou à descoberta de genes raros (de expressão restrita) que podem ser importantes do ponto de vista terapêutico para o tratamento de algumas formas de câncer”, comenta Simpson, um dos autores do estudo, que foi coordenador do Genoma Humano do Câncer, encerrado em junho deste ano, e hoje trabalha na sede internacional do Instituto Ludwig, em Nova York. Em menor escala, também foram obtidas seqüências expressas de outros tipos de tecidos, sobretudo da próstata e do ovário.

Para entender o papel dos genes no desenvolvimento dos principais tipos de câncer e em seus respectivos tecidos normais, ambas as iniciativas geraram muitas ESTs. Para ser mais preciso, o

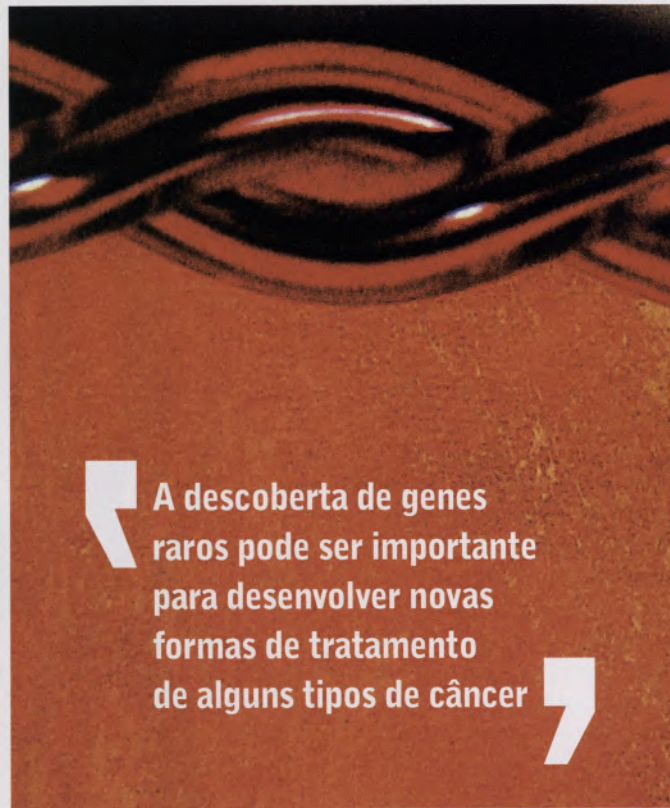
No coração dos genes

Técnica brasileira mostra quais trechos de DNA são ativados nas células sadias e nos tumores



Genoma Humano do Câncer produziu 823 mil seqüências expressas e o CGAP, 1,2 milhão de fragmentos de genes. Juntos, os dois projetos, que firmaram uma parceria há alguns anos, geraram mais de 2 milhões de ESTs. “Esse número equivale a mais ou menos 40% de todas as seqüências expressas derivadas de tecidos humanos depositadas nos bancos públicos de dados”, diz Simpson. Estima-se que os 2 milhões de seqüências expressas extraídas de tumores e as respectivas células saudáveis estejam relacionados a 23.500 genes humanos, cerca de três quartos de todos os genes conhecidos do *Homo sapiens*. Entre os tecidos estudados, as células do pulmão foram as que apresentaram o maior número de genes ativos, 13.390. “Isoladamente, nenhum tipo de tecido expressou mais do que 57% dos genes que estavam representados em nossas ESTs”, afirma Dias Neto. As células de mama se utilizaram do menor número de genes, 10.380. Nos demais tecidos, a quantidade de genes expressos foi de 10 mil a 13 mil.

Para que o leitor não fique perdido em meio a tantos números, são necessárias algumas explicações sobre genes expressos (ativos) e as metodologias utilizadas para obtenção de ESTs. A molécula de DNA de uma pessoa é idêntica em qualquer uma das células com núcleo. Portanto, qualquer tecido tem o mesmo genoma, o mesmo conjunto de genes. De acordo com as projeções mais recentes, a espécie humana apresenta cerca de 29 mil genes. Mas o fato de todos os tipos de tecidos humanos terem os mesmos genes não significa que todas as células usam esses genes de forma igual. Quando acionado, um gene despacha, com o auxílio de uma outra molécula (o RNA mensageiro), uma receita química para produzir a proteína especificamente associada a ele. Se o gene não está ativo, a proteína não é sintetizada.



A descoberta de genes raros pode ser importante para desenvolver novas formas de tratamento de alguns tipos de câncer

Como mostra o artigo do *PNAS*, em alguns tipos de tecidos, e esse é o caso do pulmão, um número maior de genes entra em funcionamento. Em outros, como nas células da mama, uma quantidade menor de genes é utilizada. Ou, em outras palavras, é expressa ou ativada. Portanto, todos os tecidos têm os mesmos genes, na mesma quantidade. Mas cada tipo de célula expressa ou ativa um subconjunto particular do número total de genes. Esse subconjunto particular de genes ativos é chamado transcriptoma de um tecido.

A quantidade de genes acionada numa célula também varia em função do tempo. Um tecido pode expressar menos genes num determinado estágio de desenvolvimento e mais genes em outro momento. “Quanto mais expres-

so for um gene, mais fácil será encontrar seus fragmentos (ESTs)”, compara Marco Antonio Zago, da Faculdade de Medicina de Ribeirão Preto da USP, coordenador do Programa Genoma Clínico do Câncer, financiado pela FAPESP. Um dos principais desafios dos pesquisadores que estudam a base genética dos tumores é entender a mudança no perfil de expressão de genes em tecidos saudáveis e em células com câncer.

Mas qual é a diferença entre o método Orestes e a técnica convencional de geração de ESTs? A técnica brasileira possibilita a obtenção de informações da parte central dos genes, onde tendem a se concentrar as suas regiões codificadoras: os trechos do gene que fornecem a receita química necessária à produção de

proteínas. Num gene humano típico, estima-se que pouco mais de 50% de sua seqüência de nucleotídeos (unidades químicas primordiais) façam parte da região codificadora. O resto da seqüência, embora importante, seria secundário. Já a técnica tradicional de obtenção de ESTs, adotada pelo CGAP, prioriza a busca de dados nas extremidades dos genes. Por focarem pontos distintos do DNA, um no meio e outro nas pontas dos genes, as duas metodologias se tornaram complementares e acabaram estimulando a parceria entre brasileiros e norte-americanos na área de câncer. “Com o Orestes, conseguimos pegar mais genes de expressão rara (que são pouco utilizados por um tecido) do que com a metodologia tradicional”, afirma Dias Neto.

Além de servirem para mapear genes ativos nos tecidos, as técnicas que trabalham com seqüências expressas se prestam a outras finalidades. No artigo do *PNAS*, os pesquisadores mostram que as ESTs também podem ser uma ferramenta útil na descoberta de mutações em genes aparentemente relacionados ao desenvolvimento de tumores. Depois de empregarem apenas a técnica Orestes para analisar as seqüências que formam 1.127 genes suspeitos de esta-

O PROJETO

Programa Genoma Humano do Câncer

COORDENADOR

ANDREW SIMPSON – Instituto Ludwig

INVESTIMENTO

US\$ 10 milhões (FAPESP) e
US\$ 10 milhões (Instituto Ludwig)

rem implicados na gênese de alguns tipos de câncer, eles foram capazes de, por exemplo, identificar 30 prováveis novos SNPs, um tipo específico de mutação. "O Orestes não é um método excelente na busca desse objetivo, mas, sem dúvida, também pode ser usado para procurar SNPs", comenta Zago. A sigla SNP – do inglês *single nucleotide polymorphism*, ou polimorfismo de um único nucleotídeo – designa as várias formas que um nucleotídeo pode assumir. Na verdade, essas possibilidades restringem-se às quatro bases nitrogenadas que formam o DNA: adenina (A), citosina (C), guanina (G) ou timina (T). Portanto, quando anunciam que descobriram um SNP relacionado a uma determinada doença, como o câncer, os cientistas estão dizendo que encontraram uma variação de apenas uma base num trecho de um gene que pode aumentar a chance de ocorrência dessa enfermidade.

Um outro ponto explorado no artigo do *PNAS* – o quarto decorrente do Programa Genoma Humano do Câncer publicado nessa revista – focaliza a utilização de seqüências expressas para determinar a ocorrência de um fenômeno conhecido como *splicing* alternativo. Quando usa de forma não tradicional as suas informações, um gene envia para a célula uma receita química ligeiramente diferente da normal. Resultado: a célula produz uma proteína distinta da que seria originalmente sintetizada. A exemplo de algumas mutações, certas formas de *splicing* alternativo podem estar relacionadas ao surgimento de doenças.

No estudo do *PNAS*, os cientistas lançaram mão de duas metodologias distintas de análise e estimaram a taxa desse fenômeno no grupo de 1.200 genes suspeitos de estarem implicados na gênese de tumores. As contas sugerem que entre 21% e 47,5% desses genes apresentam *splicing* alternativo. Agora, o desafio é saber quais dessas alterações podem ser patogênicas e quais são inócuas. "O Genoma Humano do Câncer e o CGAP tornaram públicos muitos dados sobre genes expressos em câncer", afirma Simpson. "Cabe agora aos pesquisadores dessa área trabalhar em profundidade os dados gerados por esses projetos."

PRECISÃO E CONFIABILIDADE

Em análise de DNA e RNA

A SINC, Representante exclusiva da TRANSGENOMIC no Brasil, conta com uma equipe altamente especializada, fornecendo todo suporte aos usuários do

SISTEMA WAVE®.

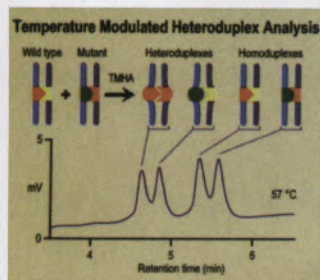
Deteção de mutações gênicas: SNPs, inserções, deleções, translocações.

Genoma humano, de plantas, animais e microrganismos.



Pesquisa e diagnóstico de doenças genéticas, hereditárias, infecciosas e câncer.

- Alta Sensibilidade • 99% de exatidão sem falsos positivos
- Reduzido tempo de análise (4-8 min) • 4000 injeções por coluna
- Totalmente automatizado • Economicamente viável



- ▶ Não há necessidade de purificação da amostra antes da análise.
- ▶ A detecção de mutação gênica é baseada na técnica de DHPLC.
- ▶ O DNA contendo polimorfismo ou mutação é hibridizado com o tipo selvagem e os fragmentos são analisados no SISTEMA WAVE®. Os *heteroduplexes* são separados dos *homoduplexes* usando uma coluna especial (DNASep®).

SINC DO BRASIL



TRANSGENOMIC

▶ **Matriz:** São Paulo - R. Cel. Melo de Oliveira, 548 - Pompéia - CEP: 05011-040
Fone (011) 3864-1411 - Fax (011) 3872-9749 E-mail: vendas@sinc.com.br - Homepage:
www.sinc.com.br <http://www.sinc.com.br>

▶ **Filiais:** • Rio de Janeiro - Fone (021) 3393-2189 - sincrj@sinc.com.br • Belo Horizonte
Fone (031) 3484-2453 - sincbh@sinc.com.br

• Florianópolis - Fone (048) 234-1483 - sincsc@sinc.com.br