

# Do Câncer ao Genoma Humano

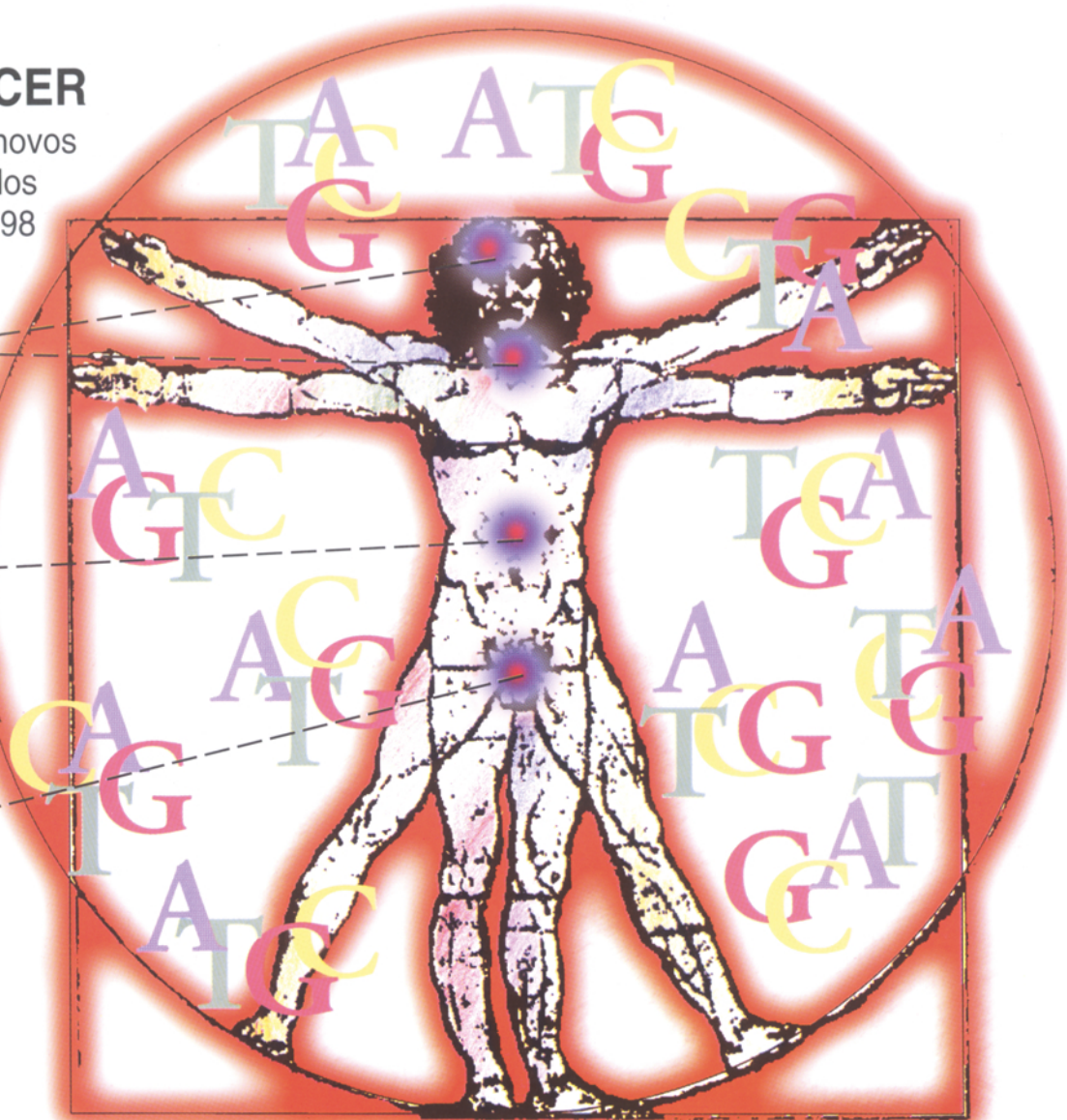
## CÂNCER

Casos novos  
estimados  
para 1998

Cabeça e  
pescoço  
20%  
(São Paulo)\*

Gástrico  
7,7%  
(Brasil)\*\*

Colo do útero  
8,1%  
(Brasil)\*\*



\* FONTE: HOSPITAL DO CÂNCER

\*\* FONTE: INSTITUTO NACIONAL DO CÂNCER

O programa Genoma/Fapesp consolida neste mês de março mais um projeto inovador entre aqueles que nasceram da idéia pioneira de seqüenciar o genoma da *Xylella fastidiosa*. Agora que o seqüenciamento da bactéria causadora da Clorose Variegada de Citros aproxima-se do fim, os esforços da rede virtual de laboratórios Onsa — Organização para Seqüenciamento e Análise de Nucleotídeos — voltam-se para o genoma da espécie humana. Através de um acordo de cooperação, a FAPESP e o Instituto Ludwig de Pesquisa sobre o Câncer vão destinar até US\$ 10 milhões nos próximos dois anos para

produzir seqüências de genes humanos e construir, com elas, uma base de dados pública. Os pesquisadores buscarão o material genético a ser analisado em tumores de cânceres os mais importantes do ponto de vista científico, no Brasil. O contrato do novo programa, Genoma Câncer, aprovado na última reunião do Conselho Superior da FAPESP, pretende gerar entre 500 mil e 750 mil seqüências num total de 200 milhões de pares de bases. As instituições prevêem que uma parte significativa dessas seqüências dará acesso a regiões codificadoras ainda não exploradas do genoma da nossa espécie.

Ambos os objetivos são expressivos, quando se considera o volume de produção de seqüências (*high throughput*) ou a proporção entre seqüências representativas de regiões codificadoras e total seqüenciado. O projeto do Nacional Institute of Cancer norte-americano, que usa estratégia semelhante à que será usada no Brasil, produziu, até o final deste mês de fevereiro, 470 mil seqüências. Por esse lado, o Genoma Câncer pode vir a representar uma contribuição efetiva dentro do esforço de elucidação genética do *Homo sapiens sapiens*, que interessa igualmente a cientistas, pesquisadores, e às indústrias farma-

cêutica, de tecnologia médica e de alimentos. Para alcançar o volume previsto de obtenção de dados, seis grandes máquinas sequenciadoras já chegaram a São Paulo. Duas delas já forneciam as primeiras seqüências na segunda quinzena de fevereiro. A capacidade delas é 10 vezes maior que a das usadas para seqüenciar a *Xylella fastidiosa*.

O aumento na capacidade de produção de seqüências, no entanto, não basta para que a FAPESP e o Instituto Ludwig cumpram a expectativa de ter até um quarto da produção representando novas regiões codificadoras. Para isso, as instituições confiam na nova estratégia de que o Instituto Ludwig já requereu a patente – segundo o contrato, “um novo conceito de seqüenciamento”.

## Conceitos

Há duas abordagens mais comuns em uso nos vários projetos de seqüenciamento desenvolvidos neste momento no mundo. Uma delas pretende o seqüenciamento completo do genoma dos organismos – como no projeto da *Xylella*. Essa estratégia funciona tanto melhor quanto menores o tamanho e a complexidade do genoma a seqüenciar. Basicamente, os pesquisadores quebram o DNA diretamente do genoma, em pedaços grandes, e os ordenam; depois partem os pedaços novamente em fragmentos de até dois mil pares de bases, que, finalmente, são clonados.

São esses fragmentos que abastecem as máquinas automáticas de seqüenciamento. As seqüências assim obtidas são, então, ordenadas segundo as sobreposições que ocorrem entre extremidades de diferentes fragmentos, localizadas através de softwares desenvolvidos especialmente para comparar cadeias de nucleotídeos. O projeto do Genoma Humano também segue esse modelo, que oferece resultados acurados e precisos.

A outra abordagem deixa de lado a idéia de seqüenciar todo o DNA de um organismo, e busca só debruçar-se sobre as regiões da cadeia de nucleotídeos que codificam proteínas. O exemplo da nossa espécie ilustra a vantagem de uma abordagem desse tipo: calcula-se que apenas cerca de 3% dos 3 bilhões de pares de bases do genoma humano formam genes, ou seja, comandam a síntese de RNA e das proteínas que regulam a vida. Essa abordagem alternativa, que não exclui a anterior, apareceu em 1991. A técnica funciona de trás para a frente: os pesquisadores partem não do DNA genômico, mas de

sua transcrição: as moléculas de RNA mensageiro. A partir delas, usando a enzima transcriptase reversa, sintetizam em laboratório fragmentos de DNA (cDNAs). O nome dado às seqüências obtidas desses fragmentos batiza o método: são as “etiquetas de seqüências expressas”; em inglês, *expressed sequence tags*, ou ESTs.

A estratégia inovadora que será posta em prática no projeto Genoma Câncer desenvolve essa abordagem, ao encontrar um caminho simples, rápido e elegante de resolver dois dos problemas até aqui inerentes a ela. O primeiro problema ocorre por causa de uma característica biológica: há genes que se expressam abundantemente, há os que se expressam moderadamente, e há os raros. A essa diferença de abundância na expressão genética correspondem moléculas de RNA mensageiro diferentemente abundantes e, por fim, também cDNAs distribuídas segundo essas irregularidades. Para resolver essa dificuldade — que resulta em trabalho desnecessário, pois muitos grupos de pesquisa acabam sintetizando e seqüenciando as mesmas ESTs, as oriundas dos genes mais abundantes —, há maneiras de realizar a “normalização” das bibliotecas de DNA obtidas a partir de RNA mensageiro. A normalização implica o uso de grandes quantidades de material genético — o que nem sempre é possível quando se trata, por exemplo, do estudo de células de tumores —, e não garante que haverá, entre os fragmentos gerados, seqüências que correspondam aos genes raros. A patente requerida pelo Instituto Ludwig encontra uma alternativa para esse problema. Outra característica do método das ESTs é gerar seqüências localizadas principalmente nas extremidades dos DNAs. O centro das regiões codificadoras (*open reading frames*) freqüentemente escapa às técnicas aplicadas até agora. A nova estratégia, chamada de ORESTES — *open reading frames ESTs* — pelo Instituto Ludwig, também contorna esse obstáculo.

Durante as negociações para a elaboração do contrato de colaboração, ambas as instituições submetem o novo conceito de seqüenciamento ao exame de especialistas e a experimentos, para obter uma avaliação prévia. Os resultados promissores justificaram a formação da parceria, a primeira entre o programa Genoma FAPESP e um instituto privado de pesquisa, sem fins lucrativos, como é o Ludwig Institute for Cancer Research. A fase piloto do projeto, fundada pelo Instituto, terminou em fevereiro

e confirmou as perspectivas favoráveis em torno da estratégia Orestes.

## A estrutura

O novo projeto apóia-se na competência científica e gerencial que o trabalho do seqüenciamento da *Xylella fastidiosa* revelou e incrementou no Estado de São Paulo, e que poderá se aprofundar com o lançamento do Genoma Câncer. A coordenação geral caberá a Andrew Simpson, hoje coordenador de DNA do projeto de seqüenciamento da bactéria da CVC. O sucesso na implantação da nova metodologia dependerá de outros três especialistas: Emmanuel Dias-Neto, responsável pela montagem das bibliotecas de Orestes; Luis Fernando Lima Reis, cuja função será garantir a pureza do RNA coletado nos tecidos tumorais e normais; e Sandro José de Souza, encarregado da bioinformática. O trabalho de seqüenciamento das Orestes caberá a cinco centros de seqüenciamento. Quatro dos coordenadores desses centros participam do Projeto *Xylella*: Sergio Verjovsky-Almeida, do Instituto de Química da USP; Marco Antonio Zago, da Fundação Hemocentro de Ribeirão Preto; Marcelo Briones, da Universidade Federal Paulista; e Maria Aparecida Nagai, da Faculdade de Medicina da USP. A eles junta-se Fernando Ferreira Costa, do Centro de Hematologia e Hemoterapia da Universidade Estadual de Campinas.

Cada um dos centros, e também o Instituto Ludwig de São Paulo, recebeu as novas máquinas automáticas. Caberá aos laboratórios de seqüenciamento alimentá-las: preparar e fornecer os clones que manterão as máquinas em operação com a máxima eficiência. A princípio, cada centro agrupará em torno de si quatro desses laboratórios. A idéia do instituto virtual, com grupos de pesquisa dispersos pelo estado e conectados por um website, outro sucesso do primeiro projeto da rede Onsa, volta a ser utilizada. Repete-se também a constituição de um Comitê Diretor, formado por especialistas reconhecidos internacionalmente: John Sgouros, que participa do Comitê do Projeto *Xylella*, do Imperial Cancer Research Fund, de Londres; Marcelo Bento-Soares, que desenvolveu a melhor técnica usada atualmente para a normalização de bibliotecas de ESTs, brasileiro, da Universidade de Iowa; e os doutores Webber Cavenee e Richard Kolodner, ambos da filial norte-americana do Instituto Ludwig e da Universidade da Califórnia, já aceitaram o convite para integrar o Comitê do Genoma Câncer.