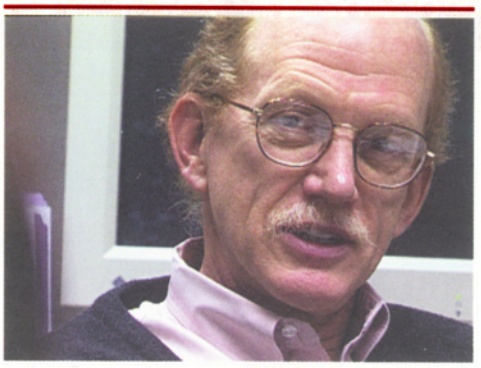
Acesso livre às informações do genoma humano

Em março passado, o vice-presidente dos Estados Unidos, Al Gore, anunciou em Washington que o governo norte-americano decidira suplementar o orçamento do Projeto Genoma Humano com mais US$ 88 milhões, para cinco laboratórios. Desse total, US$ 38 milhões – a maior fatia para um só centro de pesquisa – foi para Bob Waterston, da Washington University, em Saint Louis, Missouri. Seu centro de seqüenciamento, o quinto maior do mundo, é um importante pilar sobre o qual se apoiam o desejo e a necessidade política dos National Institutes of Health de tornar disponível, até a primavera do próximo ano, um esboço da seqüência do genoma da nossa espécie. A entrevista do doutor Waterston revela, portanto, as concepções de um cientista comprometido com os valores do establishment científico do país que mais investe em pesquisa em todo o mundo. Ele defende a publicidade imediata de todos os dados e peleja pela precisão da informação. No entendimento do doutor Waterston, o Projeto Genoma Humano oferecerá a base de dados sobre a qual os cientistas deverão se debruçar durante o próximo século; daí a importância de que os dados sejam tão confiáveis quanto possível. Expõem-se, nas duas questões, as divergências que separam o grupo financiado por fundos públicos e o ousado e inovador Craig Venter. No ano passado, com John Sulston, Waterston publicou o genoma da *C. elegans* – um nematóide de 1 milímetro de comprimento, o primeiro animal a ter seu DNA conhecido.



TERCIO PARISI

Bob Waterston

■ *Gostaria de começar pedindo ao senhor que contasse a história das decisões que envolveram o esforço de seqüenciamento da* C.elegans...

— Decidimos trabalhar para a caracterização de todo o DNA da *C. elegans* porque queríamos ter acesso a qualquer gene do nematóide a qualquer tempo. Quem começou o trabalho foi o John Sulston, agora no Sanger Center; mais tarde nos juntamos a ele. Nós tínhamos os cromossomos da *C. elegans* divididos em pequenas seqüências, que nós poderíamos organizar para ter sua representação completa. Naquele momento, 1989, o Projeto Genoma Humano decidia que outros organismos seriam estudados além do nosso. A *C. elegans* não estava na lista, e a comunidade em torno dela agitou-se: ficou claro que ter a seqüência completa traria vantagens, mesmo que isso soasse como uma fantasia naquele momento, inclusive por causa do investimento que seria gerado pela tentativa de seqüenciar. A *C. elegans* é um organismo-modelo e, se a levedura e a drosófila iriam ter seus genomas seqüenciados, então os pesquisadores estudando *C. elegans* estariam em desvantagem. A comunidade então decidiu reagir, e conseguimos marcar um encontro com James Watson, diretor do laboratório de Cold Spring Harbour. Nós dissemos a ele que tínhamos o mapa, que estávamos prontos para entrar no seqüenciamento. A pressão da comunidade levou John e eu a nos dispormos a passar de mapear o genoma para seqüenciá-lo. E a *C. elegans* acabou incluída como um entre os cinco organismos selecionados para serem seqüenciados além do humano.

■ *Além de não ficar em desvantagem, houve outros ganhos com o seqüenciamento da* C. elegans?

— Para responder sua pergunta, vou precisar elogiar um pouco a mim mesmo. O fato de John e eu termos envolvido nossos grupos nisto teve um impacto significativo, tanto na forma de seqüenciamento quanto na maneira de lidar com os dados. Somos ambos pragmáticos e queríamos realizar o trabalho, este era o nosso maior objetivo; não nos importava qual seria a tecnologia, que programa usaríamos. Isso nos libertava dos

constrangimentos que pesavam sobre outros centros de seqüenciamento; o fato de sermos novos no negócio também tornou-se vantajoso, porque pudemos escolher soluções que funcionavam. Conseguimos diminuir muito os custos naquele momento, o que nos deu a possibilidade de impulsionar outros projetos. A participação dos nossos dois grupos para completar o seqüenciamento da levedura foi central para apressar o cronograma dos grupos europeus envolvidos. Em vez de ficar pronta em 98, a levedura acabou em 95. Foi bom negócio para eles. O outro aspecto importante é que desde o começo, mesmo quando ainda estávamos no mapa, nós dividimos nossos dados com a comunidade o mais precocemente possível. Ainda não havia Internet, e não podíamos então oferecer os dados a cada 24 horas; mas encontramos uma maneira de divulgar de três em três meses, o que era essencialmente um produto não finalizado *(work in progress)*. Nós aplicamos a mesma política para os dados do seqüenciamento e, assim que foi possível divulgar diariamente os dados, passamos a fazê-lo. Ficou muito claro pela resposta dos pesquisadores da *C. elegans* que eles consideraram a disponibilidade dos dados altamente benéfica. Não reservávamos a informação para nossos próprios propósitos e para nossos próprios interesses. Essa era uma prática bastante diferente da tradição até aquele momento.

> "O cientista quer a informação tão logo ela esteja disponível, mesmo que ainda imperfeita"

▪ *O que o senhor gostaria que fosse o entendimento da extensão das patentes nesta área?*

— No que concerne aos cientistas, é claro que eles desejam os dados tão logo possam obtê-los, e incluo aí os cientistas da indústria e da academia. O cientista quer a informação assim que ela esteja disponível, mesmo que ainda imperfeita. Fomos pagos para determinar a seqüência de pedaços de DNA importantes para eles; se colocarmos à disposição a seqüência bruta, eles irão buscar a informação lá, e estarão felizes com ela, mesmo com os defeitos. Logo, a pergunta é qual é a melhor maneira de assegurar que esta informação resulte no maior avanço possível para a ciência e para a medicina. Isso nos leva a toda uma questão filosófica: as patentes restringem o avanço ou incentivam o avanço? No caso de seqüenciamento de DNA, há um diferencial, que é o volume de informação. Um centro como o nosso obtém mais informação do que jamais poderá explorar. Imagino que todo o próximo século será necessário para compreender o genoma humano, com centenas de milhares de pessoas trabalhando nele. Vai levar todo esse tempo e todo esse esforço para compreendê-lo e analisá-lo de uma maneira que faça sentido. Mas nós podemos obter a informação, no ano que vem vamos gerar um terço do genoma aqui neste centro. Teríamos o direito de controlar o desenvolvimento dessa informação e decidir quem poderá ter o privilegiado acesso a essa informação?

▪ *A política em relação à divulgação de dados preocupa-o no caso de companhias como a Celera ou a Incyte?*

— Essencialmente, essas companhias vão restringir o acesso a seus dados. Temos também de notar que eles ainda não fizeram nada e que suas políticas podem mudar, segundo as necessidades do negócio. No caso do Genoma Humano, como há muita gente interessada, e pelo fato de suas implicações serem tão vastas, acho que há razão para preocupação. Seqüenciamento vai se tornar parte da indústria, a atividade ainda é desenvolvida na academia porque há organismos que interessam à academia e não interessam à indústria. Claro que, no caso do genoma humano, todos se interessam. Sim, me preocupa a questão do acesso aos dados, e me preocupa também a precisão dos dados que estão sendo gerados. Ambas as coisas são muito importantes. Até agora, quem se provou capaz de atender aos dois requisitos foi o grupo que recebe financiamento público: na última verificação que fizemos, os grandes centros apresentavam menos de um erro a cada mil bases. Aqui na Universidade de Washington, em trechos de 2 mil bases não foi encontrado nenhum erro.

▪ *Por que os senhores decidiram publicar o paper da* C. elegans, *apesar de ainda haver buracos na seqüência?*

— Estimávamos, quando publicamos, que havia mais ou menos 135 seqüências com buracos; agora, o número já caiu para 95 ou 100, por aí – quer dizer, já houve progressos. Os buracos remanescentes são realmente muito difíceis de ser seqüenciados, e ainda não está claro se será possível fazê-lo. Também parece claro que não há genes nesses trechos com problemas, eles contêm pouca informação biológica. Há outro aspecto peculiar sobre a *C. elegans:* em vez de ela ter centrômeros confinados em um determinado ponto do cromossomo, ela tem, na verdade, centrômeros dispersos. Nos mamíferos, os centrômeros

podem ser a repetição continuada de uma mesma seqüência. Então, o que encontramos na *C. elegans* é um segmento aqui e outro ali e outro mais adiante, da mesma seqüência incrivelmente repetida. Ninguém realmente até agora afirmou que vai seqüenciar corretamente os centrômeros humanos. Todos reconhecem que é uma impossibilidade; na *C. elegans* temos o mesmo problema, distribuído por uma porção de lugares; foi isso que deixamos para trás: repetições que serão ignoradas em outros genomas. Quando, no final do ano passado, aprontamos a maior parte do que era relevante, se não tudo que era biologicamente relevante, e já que não sabíamos quanto tempo o resto do esforço iria custar, pareceu razoável ir em frente e publicar. O que estamos tentando fazer agora é descrever essas regiões o mais acuradamente possível. Sabemos que não obteremos a seqüência precisa de alguns desses lugares, porque eles são realmente muito complicados e muito resistentes às técnicas atuais. Com a descrição minuciosa, queremos ficar certos de que não estamos perdendo nada importante. Alguns desses segmentos são 30 a 50 mil bases de apenas uma simples repetição, e descobrir se não há nada nesses trechos é realmente um processo difícil. Fazemos essa verificação exaustivamente, sobre cada um dos pedaços, antes de desistir. Parece que há segmentos assim também no genoma humano, que não podem ser seqüenciados com o uso das técnicas atuais, mas os problemas são muito menos freqüentes. Enquanto na *C. elegans* encontramos um segmento difícil a cada 10 mil ou 50 mil bases, no genoma humano as regiões de dificuldade aparecem 10 ou 100 vezes mais raramente.

> "O governo foi responsável pela conceitualização do Projeto Genoma Humano e é responsável até hoje por seu progresso"

■ *Qual o achado científico mais interessante que o senhor encontrou no seqüenciamento da* C. elegans?

— Notável na *C. elegans* são os agrupamentos (clusters) de genes. Em segmentos com 10 genes, encontramos sete extremamente próximos uns dos outros e, com o que sabemos até agora, todos provavelmente fazendo a mesma coisa. Nós esperávamos encontrar isso, mas não na extensão em que de fato encontramos. Temos muito boa idéia sobre o que fazem alguns desses agrupamentos de genes, quando podemos estudá-los em outras espécies. Podemos comparar a *C. elegans* com a levedura, por exemplo. No entanto, há ainda um número expressivo de genes sobre os quais nada sabemos. Há numerosos desses agrupamentos, e os pesquisadores precisam imaginar o que eles fazem. Não é fácil porque, se você simplesmente elimina um gene entre as cópias que são encontradas no cluster, práticamente nada acontece. Será um desafio descobrir mais sobre eles. Parece claro que essa organização faz parte da estratégia que o nematóide vem usando para adaptar seu genoma ao ambiente. Relacionado também à evolução, há outro dado: os cromossomos não são uniformes em termos da taxa de mudança. Cada cromossomo parece ter uma região no centro, onde se localizam os genes mais conservados, mais antigos, que estão bem representados na levedura. Essa região central tem menos segmentos repetidos, que aparecem muito mais nas pontas dos cromossomos. Algumas vezes são seqüências repetidas que só aparecem ali e que nunca encontramos na região central. Além do mais, os genes nas pontas tendem a ser menos conservados, e também a se expressar menos abundantemente. Há uma fronteira clara, um limite bem definido, como se houvesse algum mecanismo pelo qual o DNA das pontas tem permissão para mudar, para tolerar seqüências repetidas. As mutações nesses genes parecem acontecer com mais freqüência; também é verdade que nas pontas há conjuntos de genes diferentes dos que encontramos no centro, que são mais importantes e mais conservados. Isso também se reflete na genética, porque quando você procura por uma mutação que tem efeitos maiores no fenótipo da minhoca, eles tendem a estar nas regiões centrais e menos freqüentemente nas pontas. No caso dos cromossomos humanos, não sabemos se há segmentos evoluindo a taxas diferentes, como parece ser o caso da minhoca. Não sabemos em quantos segmentos diferentes podemos dividir o cromossomo humano, mas está claro que ele tem algum tipo de segmentação, não é tão simples quanto na minhoca, porque nela há claramente uma região mais conservada flanqueada por esses segmentos mais variáveis. É bem possível que haja mecanismos na espécie humana para separar partes diversas dos cromossomos para diferentes propósitos.

■ *O senhor acaba de receber um grant de US$ 38 milhões, anunciado pelo vice-presidente da República, com a missão de apressar os trabalhos do Projeto Genoma Humano e entregar uma primeira versão da seqüência dentro de um ano. Isso será possível?*

— Em primeiro lugar, confiamos que vamos cumprir o objetivo; esta é uma decisão séria e não apenas relações

públicas. Os cinco maiores centros entre os financiados por dinheiro público já produzem alguma coisa como 14 milhões de leituras (*reads*) por ano; se nós produzirmos duas vezes ou duas vezes e meia esse número durante os próximos 12 meses, aprontaremos a primeira versão da seqüência. Nosso centro tem dobrado sua capacidade de seqüenciamento a cada ano desde que nós começamos; então, não será novo para nós. Neste momento, fazemos 4,5 milhões de leituras por ano, e nossa expectativa é chegarmos a 9 ou 10 milhões. Então, é realmente possível executar a tarefa. Mas a decisão de lançar esta primeira versão surgiu de conversas com líderes da indústria farmacêutica e membros da comunidade que estuda genética humana. Eles querem acesso à seqüência já, o mais rápido possível – e achamos que este é o melhor meio de estimular a pesquisa. O que estamos tentando fazer é usar bem esses fundos do governo.

> "Acho que o resultado final será que entenderemos o papel dos genes muito mais claramente"

▪ *Mas Haseltine, da Human Genome Sciences, declarou aos jornais que o governo deveria sair desse tipo de projeto...*

— Acho incrível essa declaração de Haseltine. Não estaríamos no ponto em que estamos hoje se o governo não tivesse se envolvido no assunto. O governo foi o responsável pela conceitualização do projeto, e é responsável por seu progresso até hoje. O Projeto Genoma Humano existe por causa do esforço de financiamento público. E, na verdade, afora a contribuição em ESTs, que não é genômica propriamente dita, os grupos privados contribuíram quase nada com seqüenciamento de genomas. Acho que há uma única bactéria cujo genoma foi feito por uma companhia privada. Parte disso deve-se ao fato de que o foco da indústria é diferente do foco da academia e do sistema público. Nós tentamos ver a longo prazo, e fornecer um produto que faça sentido e seja durável.

As ESTs são um exemplo ilustrativo das diferenças. Há muito entusiasmo na indústria pelas ESTs, o que não se dá na academia, e a razão é que as ESTs fornecem informação rápida, e malfeita ("quick and dirty"). É informação transitória, porque, quando o genoma humano terminar, ela terá muito pouco valor. As ESTs são importantes na tentativa de interpretar as seqüências, nesse sentido elas têm um valor mais durável porque são registros de substâncias expressas dentro da célula. Há objetivos diferentes. Não está claro, por exemplo, o que vai ser o produto da Celera. Os padrões que adotamos dentro do setor público do Projeto Genoma Humano não foram adotados pela Celera; não sabemos se eles serão capazes de alcançar alta qualidade de produto com o método que estão usando. Pessoalmente, acho que não.

▪ *Há muita esperança depositada sobre o seqüenciamento do genoma humano. A genômica é apresentada como um caminho para a imortalidade, há promessas de vida mais longa, promessas de botar um ponto final no sofrimento humano. Não é demais?*

— Bem, nós vivemos na era do exagero, do sensacionalismo. Em parte por causa da TV, estamos todos competindo por aparecer, e o exagero é o caminho para obter exposição; então, claro, concordo que a biotecnologia, e a genômica em particular, tem sido também objeto de exagero. E isso não vem só do setor privado, o setor acadêmico também tem culpa. O fato é que os genes e a genética estão envolvidos em quase tudo. Estão claramente envolvidos no fato de eu estar ficando careca e de eu ser ruivo, o que são coisas triviais; mas também, provavelmente, contribuem para eu ter ou não hipertensão, para ter ou não Alzheimer. Os genes provavelmente definem minhas habilidades atléticas, minha inteligência e daí para frente; eles estão por trás do que fazemos. Também há a questão do ambiente: algumas coisas são mais influenciadas pelo ambiente, outras mais pelos genes. Inteligência, por exemplo – não sabemos se se deve mais aos genes ou ao ambiente.

▪ *O senhor acredita que vai haver reparo genético para a obtenção da felicidade?*

— Penso que isso está além do que podemos esperar. Nós não entendemos a felicidade. Estamos começando agora a entender a que se deve a fibrose cística, de uma maneira que não entendíamos antes de conhecer os genes que se relacionam a essa doença. Penso que nós encontraremos genes que trazem uma disposição para o bom humor, em oposição a uma disposição para o mau humor. Mas esses genes não serão a resposta – se coisas más acontecem, provavelmente a tristeza virá. Será muito complexo, não haverá uma solução "tire este gene e tudo vai ficar bem". No entanto, acho que o resultado final será que entenderemos o papel dos genes muito mais claramente e seremos capazes de manipular a maneira como influenciam a natureza humana, a condição humana.