

# Projeto Genoma

Fernando de Castro Reinach

A biologia dedicou uma considerável fração de seu esforço produtivo na elaboração de listas. Listas de espécies animais, listas de hormônios, listas de processos metabólicos, listas de componentes celulares, e assim vai. Para os críticos dessa atitude o conhecimento horizontal e superficial teve precedência sobre estudos verticalizados. Os elaboradores de listas se defendem argüindo que as grandes sínteses, como o darwinismo, têm como matéria-prima compilações cuidadosas da diversidade horizontal apresentada pela natureza.

Essa polêmica voltou à tona com a recente decisão do governo dos EUA de investir bilhões de dólares, ao longo dos próximos trinta anos, com o objetivo de seqüenciar o genoma humano, produzindo a maior lista da história da ciência. Essa lista conterà toda a informação genética que determina o que é o homem. Esse é o primeiro megaprojeto das ciências biológicas, e seus possíveis desdobramentos têm causado celeuma. O objetivo deste artigo é explicar a natureza desse projeto e algumas de suas implicações.

Quando uma família se reúne em volta de um recém-nascido e discute se o nariz é parecido com o do pai ou se os olhos se parecem com os da mãe, estão na verdade tentando avaliar a contribuição relativa do genoma paterno e materno na constituição do recém-nascido. O genoma contém as informações passadas de pai para filho através do processo de reprodução. Assim, o genoma de cada um de nós é constituído de duas metades, uma proveniente de nosso pai (via espermatozóide), outra de nossa mãe (via óvulo). A informação presente no ovo (seu genoma) será a determinante de todas as suas características genéticas, como a forma do nariz, cor dos olhos e da pele.

É importante lembrar que nem todas as determinantes de nossas características são genéticas. Um exemplo típico é minha capacidade de andar de bicicleta. A informação necessária para esse feito veio de meus pais. Essa informação não é genética e seguramente não foi transmitida por meio de espermatozóide ou óvulo. Essa informação não faz parte de meu genoma, e portanto não será passada a meus filhos através de meus espermatozóides. Se eu quiser que eles andem de bicicleta, serão necessárias algumas lições e muitos tombos. Por outro lado uma série de características necessárias para que eles andem de bicicleta serão herdadas geneticamente, como por exemplo o fato deles terem pernas e um cérebro capaz de equilibrar o corpo numa bicicleta. Toda a nossa informação cultural, apesar de determinar uma grande parte de nossas características e de ser transmitida de pai para filho, não é genética.

Compreender os limites da contribuição genética e ambiental na determinação das características de cada indivíduo é essencial para entender que diferenças genéticas não podem ser usadas para justificar qualquer tipo de discriminação racial. Tomemos um segundo exemplo: Carl Lewis é capaz de correr cem metros em menos de dez segundos. Eu não faço o mesmo em menos de vinte segundos. Qual a razão para essa diferença? É claro que se eu tivesse me dedicado ao atletismo poderia correr cem metros em treze ou dezesseis segundos, demonstrando que parte de nossa diferença provém de estímulos

**FERNANDO DE CASTRO REINACH** é professor do Departamento de Bioquímica do Instituto de Química da USP.

ambientais. Mas será que eu poderia chegar aos dez segundos? Duvido, pois apesar de ter sido bem alimentado nunca passei de 1,70m de altura e tenho pernas curtas. Provavelmente minha constituição genética nunca me permitiria chegar aos dez segundos. Podemos concluir que Carl Lewis, para a tarefa de correr distâncias curtas, é geneticamente superior. Outras pessoas prefeririam dizer que ele tem um talento nato para o atletismo. Isso não quer dizer que ele seja geneticamente superior em qualquer parâmetro analisado, nem obrigatoriamente inferior em algum outro aspecto, mas é uma simples constatação da diferença genética existente entre os membros da espécie humana. Do ponto de vista biológico, a igualdade entre as pessoas tem uma natureza estatística. Somos todos iguais porque a média das diferenças genéticas entre as pessoas é muito menor que a influência do ambiente sobre a expressão das informações existentes no genoma. Um fato bem estabelecido pela biologia é que as características de um indivíduo (o fenótipo) é o produto de uma interação complexa entre suas características genéticas (o genótipo) e o meio ambiente. Por esse motivo, conhecer as informações presentes no genoma de um indivíduo não é suficiente para prever suas características.

A principal contribuição da biologia nesta segunda metade do século XX foi a descoberta de como o material genético presente no genoma é armazenado, transmitido de pai para filho e utilizado durante a vida. Essas descobertas equivaleriam, no campo da transmissão da informação cultural, às seguintes descobertas hipotéticas (em parênteses os equivalentes genéticos): primeiro se descobriria que uma civilização (os seres vivos) mantém um registro de toda sua informação cultural (genoma) na forma de, por exemplo, livros (moléculas de ácido desoxirribonucléico, DNA). Nestes livros, toda a informação estaria registrada em uma seqüência linear de letras (seqüências de bases nitrogenadas), das quais existiriam somente 33 tipos (quatro tipos de bases nitrogenadas: A, T, C e G). Essas letras se agrupariam em unidades chamadas capítulos (genes). Além disso, descobrir-se-ia que essa informação seria utilizada através de um processo complexo chamado de leitura (transcrição e tradução), sem o qual a informação não teria qualquer sentido. Por motivos de segurança, cada membro dessa civilização possuiria duas cópias de todos os livros (DNA). Quando nascesse um novo membro dessa civilização, tanto o pai quanto a mãe passariam ao filho uma cópia de todos os seus livros (DNAs do espermatozóide e do óvulo) de modo que o filho tivesse novamente duas cópias da informação cultural (informação genética).

O genoma de cada ser humano é constituído de 46 moléculas de DNA. Existem duas cópias de cada molécula (uma vinda da mãe, outra do pai). Portanto cada um de nós tem 23 moléculas diferentes. As moléculas de DNA são polímeros muito longos onde quatro subunidades chamadas de T, C, G e A estão arranjadas com uma ordem precisa. Toda a informação genética necessária para nossa vida está contida nesta seqüência de subunidades que são lidas e interpretadas pelas células do nosso corpo através de um mecanismo bastante complexo. Em todos os seres vivos a informação genética está contida na seqüência de subunidades das moléculas de DNA que constituem seus genomas.

O Projeto Genoma tem como objetivo conhecer a seqüência precisa destas quatro subunidades nas 23 moléculas que constituem o genoma humano. O problema de determinar a seqüência destas subunidades é conceitualmente simples uma vez que os genomas de muitos seres vivos já foram inteiramente seqüenciados. A informação obtida será nada mais que uma seqüência do tipo ATTTCCGGATTTAAGTGTCCGTATGCCAGTT onde cada letra representa uma subunidade. A dificuldade de se obter a seqüência completa do genoma humano é hoje um problema de tamanho. Os seres vivos mais simples, como os vírus, possuem genomas pequenos, da ordem de milhares de subunidades (por exemplo, o genoma do vírus da Aids, mostrado na ilustração, possui somente 9.229 subunidades e com essa informação faz o estrago que é de conhecimento de todos). Nos seres vivos mais complexos se observa um aumento do tamanho do genoma. Bactérias como *E. coli* possuem genomas bem maiores com 4.200.000 subunidades. Uma mosca possui um genoma de 140.000.000 subunidades. O genoma humano possui 3.300.000.000 subunidades. De maneira geral, o tamanho do genoma reflete a complexidade do ser vivo, pois é fácil entender que a quantidade de informação necessária para determinar as características de um ser humano é maior que a quantidade necessária para determinar as características de uma bactéria. Para se ter uma noção do tamanho do genoma humano, ou o volume de informação que será obtido no Projeto Genoma, basta lembrar que uma página de livro contém aproximadamente 2.700 letras e que um livro de 500 páginas tem 1.350.000 letras. Se ao final do Projeto Genoma fosse editada uma obra contendo toda a seqüência do genoma humano, ela teria 2.450 livros de 500 pági-

Na página ao lado, seqüência completa (9.229 bases) do DNA do vírus da Aids (HIV-1). A seqüência foi retirada do Genebank, um banco de dados internacional que coleta todas as seqüências de DNA obtidas nos diferentes laboratórios. Em sua versão 61 (setembro de 1989) o banco de dados possuía a seqüência de 35.342 moléculas de DNA, totalizando 34.762.585 bases. Atualmente, esse banco de dados cresce 15% ao ano



para um benefício muito pequeno. O argumento principal é que obteremos a seqüência e não saberemos interpretá-la. Na verdade não compreendemos ainda a maneira como a informação está contida na seqüência de bases do DNA. De posse da seqüência do genoma humano, os cientistas poderiam ser comparados a uma criança recém-alfabetizada, com um pequeno vocabulário, que fosse presenteada com as obras completas de Shakespeare. Sabemos que as moléculas de DNA codificam a informação de várias formas. A mais simples consiste na codificação da seqüência de aminoácidos das proteínas. Nessa forma a informação é lida de três em três subunidades. Cada *triplet* distinto corresponde a uma das vinte subunidades que compõem as proteínas. A tabela que relaciona cada *triplet* a um aminoácido é chamada de código genético e é idêntica para todos os seres vivos. É dessa maneira que os seres vivos utilizam as informações contidas no DNA para fabricar a hemoglobina que transporta o oxigênio em nosso sangue, as enzimas do sistema digestivo e as proteínas que constituem nosso cabelo. Porém, ainda não sabemos como a informação contida no DNA é utilizada para gerar, no feto, uma estrutura complexa como uma mão ou um cérebro. Sem dúvida, nossa capacidade de “leitura” das informações presentes no DNA é limitada. Um exemplo claro disso é que conhecemos a seqüência completa do DNA do vírus da Aids (vide figura) e nem por isso compreendemos o suficiente sobre sua biologia para podermos destruí-lo sem matar o paciente

que o carrega. Muitos críticos do projeto acreditam ser preferível investir o dinheiro em pesquisas que levassem a uma maior compreensão dos mecanismos utilizados pelos seres vivos para “ler” e “traduzir” as informações presentes no DNA. Acreditam que é prematuro obter a seqüência completa do genoma humano se ainda não podemos entendê-lo. Os defensores do projeto acreditam que se não obtivermos a seqüência será muito difícil compreender a complexidade que nos espera pela frente. Em outras palavras, os defensores do Projeto Genoma acreditam que somente praticando a leitura em obras complexas poderemos compreender toda a sutileza de Shakespeare.

Um segundo motivo de polêmica, que provoca certa reação entre a população em geral, é o problema moral do homem ter acesso a sua informação genética iniciando sua manipulação e redirecionando a seleção natural. Os mesmos medos que existiram no início da clonagem de genes em 1980 ressurgem novamente. Estes medos têm pouco cabimento. É difícil imaginar que a posse da seqüência do genoma humano permita manipulações não possíveis hoje. Os métodos clássicos de seleção, como os utilizados para melhorar as raças de bovinos, ainda são eficientes. O que barra a espécie humana de utilizar esses métodos não é o desconhecimento de como fazê-lo, mas um impedimento de ordem moral. A prática tem mostrado que todos os esforços científicos dirigidos à compreensão dos fenômenos relacionados ao genoma humano têm gerado avanços na área médica. Exemplos como a descoberta dos oncogenes (genes responsáveis pelo desencadeamento dos processos cancerosos), ou de genes responsáveis por muitas doenças genéticas, exemplificam essa argumentação.

A realidade é que o Projeto Genoma já foi iniciado. E, como muitos projetos científicos, não nos permite dizer com certeza o que iremos encontrar pela frente, nem prever as conseqüências sociais da informação obtida.

**Na verdade não entendemos  
ainda a maneira como a  
informação está contida na  
seqüência de bases do DNA.  
De posse da seqüência do  
genoma humano, os cientistas  
podem se comparar a uma  
criança recém-alfabetizada,  
com um pequeno vocabulário,  
que fosse presenteada com  
as obras completas de  
Shakespeare... Um exemplo  
claro disso é o fato de que  
conhecemos toda a seqüência  
do DNA do vírus da Aids  
e não sabemos o bastante de  
sua biologia para destruí-lo  
sem matar o paciente**

Agradeço a Ronaldo Bento Quaggio, Mário Gustavo Mayer e Elisa Grinspum a leitura crítica do manuscrito.