



Um ecossistema federado para compartilhamento de dados genômicos

18 de agosto de 2016

Karina Toledo | Agência

FAPESP – A importância de se compartilhar dados de estudos genômicos para fazer avançar o tratamento de doenças foi tema de um comentário publicado (<http://science.sciencemag.org/content/352/6291/1278>) na revista *Science*, em junho, pela equipe do consórcio The Global Alliance for Genomics and Health ([GA4GH](#)).

No texto, um dos exemplos positivos mencionados foi a Brazilian Initiative on Precision Medicine (BIPMed) – iniciativa que reúne integrante de cinco Centros de Pesquisa, Inovação e Difusão (CEPIDs) apoiados pela FAPESP e que lançou, em novembro de 2015, o [primeiro banco público](#) de dados genômicos da América Latina.

Segundo os autores do texto divulgado na *Science*, embora tenha havido progresso nos últimos anos, a maior parte dos dados clínicos e genômicos ainda são coletados e estudados em silos – compartimentados por doença, por instituição, por país.

“Os primeiros esforços de compartilhamento têm permitido desenvolver tratamento para doenças raras e algumas formas de câncer. Porém, tal benefício só atingirá toda a população quando médicos e pesquisadores puderem acessar e comparar dados de milhões de indivíduos”, escreveu a equipe da GA4GH.

O consórcio internacional foi fundado em 2013 com esse propósito: ajudar pacientes, pesquisadores e médicos a obter avanços científicos por meio do compartilhamento responsável de dados. Um dos projetos mais conhecidos do grupo é o Beacon, no qual as instituições associadas – entre elas a BIPMed – oferecem um serviço online capaz de responder perguntas simples (como respostas apenas “sim” ou “não”) e que não violem informações consideradas sigilosas. São consultas como: “Você tem em seu banco algum genoma com um 'A' na posição 100.735 no cromossomo 3?”

Os obstáculos

O compartilhamento de dados genômicos – muitas vezes protegidos por cláusulas de confidencialidade – envolve uma série de desafios éticos, logísticos e de segurança que precisam ser considerados. A invasão da privacidade dos doadores é sempre um risco. Há também o medo dos pesquisadores de perderem vantagens relacionadas ao acesso exclusivo às informações. Além disso, reunir dados de todo o mundo em um único grande banco exigiria muitos recursos e infraestrutura.

A proposta apresentada pela GA4GH no texto é a criação de um “ecossistema federado” para o compartilhamento de informações, ou seja, um tipo de parceria na qual cada instituição permanece responsável por alimentar, armazenar e gerir seu próprio banco de dados. Esses vários bancos, porém, seriam conectados por meio de um software e cada instituição teria liberdade para escolher o nível de informação que deseja compartilhar.

A proposta despertou a atenção da revista *Nature*, que publicou um [editorial](#) a respeito.

“Imagine a dificuldade de encontrar um livro específico se o conteúdo de uma dúzia de diferentes bibliotecas nacionais fosse reunido todo em um único local e, em seguida, conceber uma forma de integrar as várias maneiras com que os diferentes conteúdos são arquivados, rastreados, gravados e disponibilizados. Seria muito mais fácil pedir a cada biblioteca que guardasse seus próprios livros e que compartilhasse a informação de como encontrá-los em cada biblioteca. E se o compartilhamento de dados de ciência pudesse seguir o mesmo caminho?”, perguntam os editores da *Nature*.

Como exemplo de que isso é possível, eles citam o *Matchmaker Exchange* – outra ferramenta implementada pela GA4GH com o objetivo de ajudar a encontrar as causas genéticas de doenças raras por meio da comparação de perfis fenotípicos e genotípicos similares.

O texto da *Nature* mais uma vez chama a atenção para o risco de invasão de privacidade. Relata que estudos genéticos massivos em países como Estados Unidos, Catar, Arábia Saudita e Brasil (no caso a BIPMed) estão coletando dados genéticos de milhões de pessoas e, portanto, há chance de a identidade de um indivíduo ser descoberta por meio da mineração desses dados.

Iscia Lopes-Cendes, professora da Faculdade de Ciências Médicas da Universidade Estadual de Campinas (FCM-Unicamp) e uma das idealizadoras da BIPMed, contou que o banco de dados genômicos criado pela organização conta com ferramentas para proteger o sigilo das informações.

Uma delas, conhecida como “bootstrap”, visa impedir que programas de computador do tipo robô façam milhares de consultas sucessivas e cujo conjunto dos resultados possibilite identificar um dos genomas contidos no banco. “A cada pergunta feita são acrescentados 3 milissegundos no tempo de resposta, até que em um determinado ponto o sistema trava”,

contou.

Além disso, há dois níveis de acesso às informações. Os dados sobre variantes genômicas e transcritos presentes na população estudada são de livre consulta. Mas para acessar o exoma (conjunto de genes codificadores de proteínas) de um grupo de indivíduos, por exemplo, será preciso assinar um contrato comprometendo-se a não dividir as informações com terceiros e não disponibilizá-las publicamente. Inicialmente, isso será possível apenas para projetos de pesquisa.

“É preciso mudar a cultura acadêmica, amenizar o sentimento de posse que os pesquisadores têm com os dados. Na maioria das vezes eles são fruto de pesquisas financiadas com dinheiro público e contam com material doado por outras pessoas. Os pesquisadores financiados pelo NIH, por exemplo, têm a obrigação de tornar seus resultados públicos, sob pena de não serem mais financiados. A mesma política poderia ser adotada pelas agências de fomento do Brasil”, comentou Cendes.

Entre os CEPIDs que fazem parte da BIPMed estão o Instituto de Pesquisa sobre Neurociências e Neurotecnologia (BRAINN, na sigla em inglês), o Centro de Pesquisa em Engenharia e Ciências Computacionais (CCES, na sigla em inglês), o Centro de Pesquisa em Terapia Celular (CTC), o Centro de Pesquisa em Obesidade e Comorbidades (OCRC, na sigla em inglês) e o Centro de Pesquisa em Doenças Inflamatórias (CRID, na sigla em inglês).

Leia mais em: agencia.fapesp.br/22255/.