



## NOTA TÉCNICA Nº 06

### A Nova Era Genômica e a Biodiversidade Brasileira

Alexandre Lima Nepomuceno, LABEX US Biotecnologia  
Ricardo Elesbão Alves, LABEX US Biodiversidade  
Isabel Rodrigues Gerhardt, Embrapa Florestas  
Ricardo Augusto Dante, Embrapa Agropecuária Oeste  
Carlos Eduardo Lazarini da Fonseca, LABEX US Coordenação

Milhões de anos de evolução da biodiversidade possibilitaram o surgimento de mecanismos moleculares, bioquímicos e fisiológicos que, em combinação, foram responsáveis pela adaptação das espécies aos seus ecossistemas. O Brasil detém de 15 a 25% dessa biodiversidade, fonte potencial para a geração de novos materiais genéticos, capazes de atender a demanda crescente da sociedade moderna por novos produtos na agricultura, medicina e indústria. Até recentemente, as ferramentas disponíveis para explorar essa riqueza eram caras e ineficientes. No entanto, uma nova geração de equipamentos e técnicas genômicas tem permitido conhecer não só a sequência de DNA dos organismos, como também detalhes de expressão gênica, possibilitando entender como uma miríade de genes é regulada de forma coordenada, produzindo proteínas e metabólitos que, no final, são responsáveis pelas características das espécies que são de interesse do homem.

A primeira planta a ser sequenciada (*Arabidopsis thaliana*) levou aproximadamente 10 anos para ter o primeiro rascunho de seu genoma apresentado. Hoje, com a utilização das segunda e terceira gerações de sequenciadores (e.g. 454, da Roche; SOLiD, da Applied Biosystems; *Genome Analyzer IIe*, da Illumina) e poderosos programas de bioinformática e modelagem computacional, os genomas podem ser sequenciados, montados e relacionados a características fenotípicas específicas de cada genótipo em poucas semanas. Uma quarta geração de sequenciadores (e.g. Oxford Nanopore, Pacific Biosciences, Ion Torrent) deve permitir, em poucas horas, ou até minutos, sequenciar o genoma completo de uma espécie a preços razoáveis que podem potencialmente chegar a menos de US\$ 100/genoma.

O volume cada vez maior de dados sendo gerados nesta nova era genômica tem levado à evolução de ferramentas de bioinformática, procedimentos computacionais e sistemas de hardware mais poderosos que começam a tirar proveito do sequenciamento de alto desempenho. Em conjunto com novas metodologias de mapeamento físico, como o ótico, estes avanços irão facilitar ainda mais a reconstrução de cromossomos quando comparados a métodos tradicionais, como o mapeamento físico ou genético.

Métodos como o "*Trinity: method for de novo assembly of full-length transcripts without reference genome*", desenvolvido pelo MIT (*Massachusetts Institute of Technology*) e instituições parceiras, permitem que sequências completas de mRNA obtidas do sequenciamento massivo de cDNA sejam remontadas, habilitando o estudo profundo de transcriptomas inteiros sem a necessidade da existência de um genoma de referência. Adicionalmente, estratégias para determinação do epigenoma – como imunoprecipitação de histonas modificadas seguida de sequenciamento (*ChIP-Seq*) e sequenciamento após tratamento com bissulfito (*Bisulfite-Seq*) para identificação de bases metiladas, entre outras – permitem investigar a dinâmica de modificações epigenéticas em escala global e suas interações com condições ambientais e perturbações genéticas diversas.

Devido a tecnologias de alto desempenho, a velocidade de obtenção de conhecimento e o seu uso no desenvolvimento de produtos têm se tornado extraordinariamente rápidos. Com a adoção de novas estratégias de transformação de plantas, como o uso de *Zinc Fingers Nucleases (ZFN)* ou de *Transcription Activator-Like effectors (TALEs)*, com o contínuo desenvolvimento de espécies-modelo, com o uso de genótipos de porte reduzido que aceleram testes *in vivo* e com o aprimoramento da fenotipagem em larga escala (e.g. *Lemnatec phenotyping system*), novos fundamentos serão desenvolvidos de forma eficiente, encurtando ainda mais o tempo até a sua aplicação em soluções e novos produtos.

Este novo período da era do conhecimento genômico tem sido a base para uma nova disciplina que está emergindo na genética, denominada de biologia sintética. Em maio de 2010, pesquisadores do *J. Craig Venter Institute (JCVI)* publicaram resultados descrevendo o sucesso na construção da primeira bactéria sintética que se autorreplica. Eles sintetizaram os 1.8 milhões de pares de base cromossômicos do genoma modificado de *Mycoplasma mycoides*. A célula sintética (*M. mycoides JCVI-syn1.0*) é a prova do princípio de que genomas podem ser desenhados no computador, quimicamente sintetizados em laboratório e transplantados para uma célula recipiente para produzir uma nova célula autorreplicante controlada por um genoma sintético. Embora muitas preocupações éticas e sociais possam ser levantadas sobre esta tecnologia, suas aplicações práticas já começam a aparecer, como na produção de biodiesel a partir da cana-de-açúcar por engenharia metabólica via biologia sintética, já em teste no Brasil pela empresa *Amyris*. Certamente, nas próximas décadas, o agronegócio será transformado por essa nova tecnologia.

A China tem colocado um grande esforço científico, tecnológico e financeiro na formação de bancos de caracteres genéticos. O *Beijing Genome Institute (BGI)*, em seu projeto de sequenciamento de 1000 espécies de referência, já completou o sequenciamento e montagem do genoma de 95 espécies, sendo que outras 505 estão com seus genomas em fase final de sequenciamento e/ou montagem. Reconhecida como talvez a maior plataforma de genômica e bioinformática do mundo, com capacidade de sequenciar até 30.000 genomas humanos por ano, cientistas do *BGI* quebraram vários paradigmas desenvolvendo tecnologias que hoje em uso eram tidas como impossíveis de funcionar.

Utilizar essas novas ferramentas para caracterizar os bancos de germoplasma de nosso país, visando criar bancos de caracteres que possam ser explorados por nossos pesquisadores, é uma alternativa importante e que deve ser amplamente entendida e institucionalizada. Para espécies de valor econômico, isso já vem sendo feito pelas grandes empresas privadas em escala bilionária de investimentos. Por outro lado, focar somente em caracteres presentes em espécies comerciais limita a informação genética obtida. Buscar na biodiversidade brasileira a fonte de novos genes e estratégias de adaptação pode ajudar a garantir nossa independência e ser um grande trunfo na busca de parcerias público-privadas para desenvolvimento de produtos agrícolas de alto valor agregado.

Uma 'Plataforma de Genômica' em espécies nativas brasileiras deverá ter como alvo principalmente características de interesse direto para a agricultura, indústria e medicina. A participação da Embrapa deveria ser concentrada em recursos genéticos de alto valor agregado, produtos pré-tecnológicos e tecnológicos, em *know-how* tecnológico e em bancos de dados e de informação de espécies da biodiversidade brasileira, capazes de contribuir com caracteres envolvidos na adaptação a estresses bióticos e abióticos, assim como caracteres envolvidos na melhora da qualidade nutricional de alimentos ou que agreguem valor aos produtos agrícolas. Por exemplo, em vista dos cenários de mudanças climáticas, deveríamos buscar nas espécies adaptadas às regiões do Cerrado e Caatinga a identificação de genes e rotas metabólicas candidatas e engenheirá-las em genótipos-elite de espécies comerciais, visando à adaptação e à estabilidade produtiva em condições de estresse. Para que estas estratégias sejam rapidamente testadas, deveremos ser capazes de obter e fenotipar em larga escala os genótipos transformados, caracterizando-os agronomicamente em condições reais de campo e em ambientes que simulem cenários previstos de mudanças climáticas.

Somente com uma poderosa estrutura de biologia de sistemas, ferramentas de bioinformática e de algoritmos computacionais que integrem todos os dados gerados nos diferentes níveis de organização biológica, permitindo a comparação de múltiplos genótipos caracterizados genômica e fenotipicamente, será possível identificar variabilidade nas sequências gênicas ou regulatórias que possam estar relacionadas a mecanismos de adaptação.

A importância do agronegócio na economia brasileira é inegável. Somos um dos maiores produtores de alimentos, agroenergia e outros produtos relacionados à agricultura do mundo, e um dos poucos que poderia aumentar consideravelmente sua produção nas próximas décadas sem destruir ambientes nativos ou preservados. Entretanto, o germoplasma utilizado nas variedades brasileiras depende cada vez mais de tecnologias proprietárias que não pertencem a empresas públicas ou privadas brasileiras. Apesar da salutar abertura do mercado brasileiro de sementes, é estratégico que nossas empresas mantenham a sua presença no mercado, auxiliando na manutenção do equilíbrio de preços, mas também trazendo ao produtor e à sociedade brasileira tecnologias que inicialmente não seriam desenvolvidas por empresas de capital externo.

Quando olhamos a dependência da agricultura brasileira em insumos como fertilizantes e defensivos agrícolas, e, principalmente, as simulações que mostram como as mudanças climáticas globais poderão afetar nossa agricultura, é óbvia a urgência de criarmos recursos genéticos com base na nossa biodiversidade, visando à disponibilidade de ferramentas para inspirar nossos cientistas a desenvolver soluções criativas para os desafios à frente. Nossos pesquisadores precisam rapidamente adquirir e utilizar este conhecimento para prover maneiras inovadoras de modificar processos moleculares, celulares e fisiológicos vegetais, adicionar novas características e controlar a expressão gênica a fim de desenvolver germoplasma com características próprias de maior produtividade, tolerância a estresses, qualidade nutricional, entre outras características de interesse. Intrinsecamente multigênicas e complexas, somente através da compreensão da estrutura, da dinâmica e do controle das etapas bioquímicas e dos processos de desenvolvimento de tais características é que poderemos avançar no uso de aplicações práticas tanto no pré-melhoramento como no melhoramento genético, seja via engenharia genética, seja via seleção genômica ampla assistida por marcadores moleculares.

A Embrapa, em conjunto com outras instituições de pesquisa no Brasil, teve e continua tendo um papel chave no desenvolvimento do nosso agronegócio, quebrando vários paradigmas como, por exemplo, a produção de grãos no cerrado. O uso prático, no nosso agronegócio, da informação genética contida na biodiversidade brasileira também é um paradigma a ser quebrado. A rápida evolução das tecnologias na genômica nos coloca em um momento único na história da humanidade. Ou somos capazes de nos organizarmos com inteligência, otimizando recursos humanos e financeiros, tendo um foco muito claro como podemos melhor utilizar esta riqueza, ou outros o farão. Talvez seja este o momento de montarmos no Brasil uma proposta que agrupe as ideias discutidas acima e obtenha de nossa biodiversidade e bancos de germoplasma nossos melhores recursos para as próximas décadas. Iniciativas como a do Laboratório Multiusuário de Bioinformática da Embrapa podem ser o cerne para projetos amplos que explorem nossa biodiversidade do ponto de vista genético.

A presença do Labex nos EUA pode auxiliar na formação de redes de pesquisas e na busca e ampliação de parcerias institucionais com centros de renome na área genômica, como por exemplo, o DOE JGI (*U.S. Dept. of Energy – Joint Genome Institute*), o TGI (*The Genome Institute*) ou o JCVI (*J. Craig Venter Institute*), nos setores público, acadêmico e privado, respectivamente. Estas parcerias podem ser estabelecidas no âmbito de elaboração de projeto de interesse comum ou mesmo de treinamentos de pesquisadores, principalmente nas áreas de bioinformática e modelagem computacional.

O principal parceiro da Embrapa nos EUA, o ARS/USDA, participa de vários projetos de genômica, e o Labex pode auxiliar na identificação de tópicos de interesse comum e na formação de grupos de pesquisa que explorem a biodiversidade de ambos os países utilizando genômica de última geração. Entretanto, o ARS/USDA executa projetos complexos de genômica em conjunto com o DOE JGI, evitando duplicação de esforços e maximizando o uso de recursos públicos. Não temos no Brasil um instituto de pesquisa cuja missão principal se assemelha a do DOE JGI ou mesmo do BGI. Entretanto, se quisermos evoluir cientificamente e criar massa crítica condizente com nosso potencial científico, econômico e social, temos que começar agora a usar rapidamente *cutting-edge technologies* com foco muito claro em nossos objetivos finais. Apesar do ascendente *know-how* brasileiro em genômica, ainda precisamos de mais infraestrutura, mais financiamento e mais massa crítica. Talvez seja o momento de considerarmos a criação de redes de genômica voltadas aos nossos interesses como nação, e a Embrapa, como líder em pesquisa agronômica do país, tem um papel crucial nesse processo.