Verde boa nova

Abordagens genômicas podem contribuir para o melhoramento e a conservação de florestas

MARCELO CARNIER DORNELAS

As estratégias de melhoramento genético baseadas em marcadores moleculares foram usadas para acelerar o processo de introdução de genes em germoplasmas altamente produtivos. Mais recentemente, estratégias baseadas em ferramentas genômicas para a descoberta de novos genes, associadas a processos eficientes de transformação genética e à possibilidade de automação e miniaturização da maioria dos ensaios analíticos e funcionais, aceleraram a identificação de produtos candidatos.

As empresas florestais têm apresentado interesse na utilização dessas tecnologias para a geração, a longo prazo, da *árvore ideal*. Por outro lado, a conservação dos recursos naturais, principalmente os da biosfera das matas nativas nacionais, também pode lançar mão das abordagens genômicas como ferramenta adicional para a análise e manutenção do patrimônio genético e da biodiversidade.

Genômica em plantas

Nos últimos 10 anos foram descobertos níveis inesperados da conservação do conteúdo e da ordem de genes, durante os milhões de anos de evolução que separam gramíneas, crucíferas, leguminosas, coníferas e solanáceas.

A conservação da ordem dos genes nos cromossomos – chamada colinearidade – tem sido usada para isolar genes, localizados precisamente em mapas genéticos de espécies com genomas grandes, baseando-se em sua posição relativa em mapas gerados em espécies modelo com genomas pequenos.

Um exemplo é a localização de genes de trigo com base em mapas gerados em arroz. O alinhamento comparativo de genes, controlando características qualitativas e quantitativas entre espécies, mostra claramente que é possível localizar tanto genes principais como os QTLs – *Quantitative Trait Loci* –, responsáveis por características essencialmente quantitativas, como produção, utilização efetiva de nutrientes, resistência a estresse e rendimento em campo.

Adicionalmente, se o gene responsável pela característica em uma das espécies já foi isolado, pode ser obtido com certa facilidade um candidato imediato para o gene ou QTL correspondente. Os genes candidatos, identificados por fenótipos mutantes em uma espécie, podem ser examinados quanto a seus efeitos em características equivalentes em outras espécies.

Avalanche de dados

O valor da *Arabidopsis* (planta herbácea da família das crucíferas, a mesma da mostarda e da colza) e do arroz, como modelos para dicotiledôneas e monocotiledôneas respectivamente, tem se fortalecido pela abordagem genômica comparativa.

O progresso da bioinformática é ansiosamente esperado, uma vez que o uso dessa técnica é o único recurso que os geneticistas e melhoristas podem utilizar para lidar com a avalanche de dados genômicos gerados atualmente. Mas a bioinformática de análise comparativa interespecífica de genes – ou genômica comparativa – ainda engatinha.

Na verdade, a criação de bancos de dados anotados e concentrados em uma cultura ou uma certa característica de interesse agroflorestal poderia trazer grandes avanços.

Seqüenciamento rápido

No início do ano 2000, o GenBank – National Center for Biotechnology Information, NCBI, http://www.ncbi.nlm.nih.gov – continha aproximadamente 350 mil seqüências de plantas em seu banco de dados dbEST, sendo que o maior número dessas seqüências era de *Arabidopsis* (72.354), tomate (51.932) e arroz (47.449).

Um dos primeiros usos do seqüenciamento rápido em larga escala de cDNAs (o DNA complementar ao RNA mensageiro,

Glossário

AFLP: *do inglês Amplified Fragment Lenght Polymorphism*, ou seja, Polimorfismo do Comprimento de Fragmentos Amplificados. É um marcador molecular cuja técnica de obtenção é parecida com a do RFLP, mas antes de digerir o DNA com enzimas, ele é amplificado pela técnica da PCR. É como uma combinação do RFLP e do RAPD.

Arabidopsis: uma planta daninha muito comum na Europa, mas que não ocorre no Brasil. Há quem a chame de arabete, que vem do nome francês.

DNA repetitivo: seqüências altamente repetitivas das bases do DNA, geralmente presente entre os genes, e cuja função no genoma ainda é desconhecida. A quantidade destas repetições é diferente de organismo para organismo e pode ser usada em comparações.

Fenótipo: característica de um indivíduo, determinada pelo seu genótipo e pelas condições ambientais.

Floema: o mesmo que líber; o tecido condutor da seiva elaborada ou orgânica nos vegetais vasculares. Localiza-se para fora do lenho.

Gameta: célula sexuada dos seres vivos com número de cromossomos típico dos gametas normais, encarregada da reprodução mediante fecundação ou fusão nuclear.

Genoma: a informação genética global, necessária para a existência de um organismo. Os animais superiores têm dois genomas: um maior, no núcleo da célula, e um menor, nas organelas celulares denominadas mitocôndrias. As plantas possuem três genomas: um maior, no núcleo da célula, e dois menores: um nas mitocôndrias e um nos cloroplastos. Todo genoma é feito de DNA (ácido desoxirribonucleico), à exceção de alguns vírus, que são feitos de RNA (ácido ribonucleico).

Genótipo: o conjunto dos genes de um indivíduo.

Germoplasma ou plasma germinativo: a fonte de recursos genéticos de uma espécie ou população de organismos.

Lignina: substância que se deposita nas paredes das células vegetais, conferindo a estas notável rigidez. É o que dá consistência à madeira, a qual pode conter até 25% de lignina.

Mapa genético saturado: um mapa do genoma, contendo marcadores genéticos a menos de um centimorgan um dos outro. O centimorgan é a unidade de medida de recombinação genética. É a unidade usada para calcular as distâncias entre os genes dentro do genoma. Os mapas genéticos têm as distâncias calculadas em centimorgans (distância relativa) ou em pares de bases do DNA (distância real).

Microssatélite: um tipo de marcador molecular que baseia-se nos diferentes tamanhos de fragmentos de DNA de seqüência repetitiva.

Nucleotídeo: unidade constituinte dos ácidos ribonucleicos e desoxirribonucleicos que, por hidrólise, fornece um açúcar, ácido fosfórico e uma base nitrogenada.

PCR: do inglês *Polymerase Chain Reaction*, ou seja, Reação em Cadeia da Polimerase. É uma técnica de biologia molecular que permite o aumento do número de cópias de uma seqüência de DNA artificialmente, fora do corpo do organismo, por uma reação química em cadeia.

Produto candidato: uma proteína ou outra substância produzida pela ação de um ou mais genes, oriunda do metabolismo da planta, interessante para o uso agroflorestal.

RAPD: do inglês *Random Amplified Polimorphic DNA*, ou seja, Polimorfismo de DNA Amplificado ao Acaso. É um tipo de marcador molecular que se baseia nos diferentes tamanhos de fragmentos de DNA (repetitivos ou não), obtidos pela técnica da PCR.

RFLP: do inglês *Restriction Fragment Lenght Polymorphism*, ou seja, Polimorfismo do Comprimento de Fragmentos de Restrição. É um tipo de marcador molecular que se baseia nos diferentes tamanhos de fragmentos de DNA não-repetitivo obtidos pelo tratamento químico do DNA com enzimas que o cortam em fragmentos de diferentes tamanhos, dependendo de sua composição de bases. Assim, organismos de composições genéticas diferentes apresentam padrões de digestão enzimática diferentes.

Xilema: o mesmo que lenho; o principal tecido de sustentação e condução da seiva bruta nos caules e raízes, caracterizado pela presença de elementos traqueais ou tubos.

ou seja, o produto da expressão de um gene) ou coleções de etiquetas de seqüências expressas – os denominados ESTs, *Expressed Sequence Tags* –, foi a identificação de genes envolvidos em passos metabólicos específicos de plantas.

As plantas sintetizam mais de 100 mil compostos diferentes, muitos dos quais com importante valor nutricional, médico ou industrial. Esses compostos valiosos são freqüentemente encontrados em espécies exóticas e não em espécies comumente cultivadas, cujos genomas têm sido alvo de intensa pesquisa.

Via de isolamento

Uma vez que interesse introduzir determinados genes numa planta cultivada, seja por introgressão genética tradicional, seja por engenharia genética, são necessários métodos eficientes para identificação e isolamento desses genes.

No passado, era necessário isolar as enzimas-alvo de tecidos exóticos para obter a informação molecular necessária à clonagem de seu gene codificador. Era um esforço geralmente demorado, caro e, não raro, malsucedido.

Recentemente, o seqüenciamento em larga escala de cDNAs preparados de tecidos com atividades metabólicas específicas fornece uma via alternativa rápida e de baixo custo efetivo para o isolamento de vários genes.

Melhoramento genético

Atualmente, mais da metade dos ESTs pode ser identificada pela busca de homologias com outras seqüências nucleotídicas ou com seqüências protéicas deduzidas dessas seqüências. Têm sido estabelecidos programas de genômica funcional para auxiliar a identificação dos genes restantes.

O alinhamento comparativo de

genes mostra claramente que é possível localizar tanto genes principais como os QTLs – Quantitative Trait Loci –, responsáveis por características quantitativas: produção, utilização efetiva de nutrientes, resistência a estresse e rendimento em campo.

Num grande número de casos, os genes identificados foram contemplados com projetos de melhoramento genético. Esses programas incluem análises de expressão com auxílio de técnicas como o *microarray*, comparações de bibliotecas de DNA e procedimentos específicos de seqüenciamento, que permitem o acompanhamento das modificações nos níveis de expressão de certos genes, segundo a aplicação de estresses bióticos, abióticos ou químicos.

Ferramentas de análise

Novas ferramentas de análises de proteínas, como a identificação de proteínas expressas diferencialmente por eletroforese bidimensional seguida de análise de espectrometria de massa, também estão sendo utilizadas na identificação da função de novos genes em *Arabidopsis*.

Ancorados geneticamente, mapas físicos de plantas cultivadas e estrategicamente importantes estão sendo desenvolvidos para ser usados como esqueletos na localização de ESTs de segmentos particulares de um dado genoma, facilitando a identificação e a clonagem de genes, baseadas em suas posições. Dessa forma, um bom mapa genético pode, em alguns casos, ser tão vantajoso quanto um projeto genoma, do ponto de vista do melhoramento.

Genômica e florestas

Vários projetos de seqüenciamento genômico têm se concentrado na identificação dos genes envolvidos na formação da madeira em gimnospermas e angiospermas.

O avanço de alguns desses projetos pode ser acompanhado atualmente pela Internet, através de um portal de informações



em genômica, o Núcleo de Pesquisas Genômicas em Espécies Florestais (NPGEF) mantido pelo Instituto de Pesquisas Florestais, IPEF. O endereço do *site* é http://www.ipef.br/pesquisa/genoma/.

Esses projetos são realizados principalmente por universidades estrangeiras e por *joint-ventures* de empresas multinacionais da área de papel e celulose.

Projetos tratam de madeira

Um interessante projeto de seqüenciamento em larga escala foi realizados nos Estados Unidos. Combinaram-se 1.097 ESTs de quatro bibliotecas em um único conjunto de dados. As quatro bibliotecas foram derivadas de árvores de pínus, especialmente de madeira de compressão, madeira normal e duas outras bibliotecas reciprocamente subtraídas, derivadas das duas primeiras.

Aproximadamente 10% dos ESTs codificavam proteínas similares a proteínas envolvidas com a síntese de carboidratos ou de ligninas e proteínas associadas à formação da parede celular, sugerindo que esse conjunto de seqüências geradas está enriquecido com genes que codificam proteínas envolvidas na biossíntese de celulose.

Um segundo projeto, ainda nos EUA, visava também a identificação de genes relacionados à formação da madeira. Envolvia duas espécies de choupo e duas bibliotecas diferentes: a mais geral, da região do câmbio de uma das espécies, supostamente enriquecida com seqüências envolvidas na formação do xilema e do floema, e uma biblioteca de cDNA mais específica, derivada da região de formação do xilema da outra espécie.

Projetos trazem evidências e sugestões

Foram obtidos ao todo 5.692 ESTs, 4% dos quais relacionados com a formação da madeira. Observou-se que os ESTs codificando proteínas envolvidas com a síntese do xilema eram duas vezes mais comuns entre os 883 ESTs derivados da biblioteca do xilema em desenvolvimento do que entre os da biblioteca mais geral, da região do câmbio.

Os conjuntos de dados derivados de pínus e de choupo eram de composição similar, no que diz respeito à freqüência de ESTs codificando proteínas envolvidas na formação de madeira e as de função desconhecida.

Isso evidencia a conservação de genes entre espécies bastante distantes geneticamente, e sugere que a genômica comparativa de árvores pode ser uma ferramenta poderosa na identificação de genes de interesse.

Genes envolvidos na xilogênese e no processo de lignificação também foram estudados em espécies herbáceas modelo, como *Arabidopsis thaliana*, e os homólogos dos genes descobertos estão sendo clonados em eucaliptos e pínus.

Um longo processo de seleção

O melhoramento genético de árvores é limitado pelo período de tempo necessário para o alcance da maturidade sexual e pelo tempo de espera necessário à avaliação de determinadas características, como a *performance* de crescimento.

Adicionalmente, as avaliações de interações entre genótipo e ambiente requerem grandes investimentos em termos de tempo e espaço físico. Nesse contexto, qualquer ferramenta capaz de auxiliar o processo de seleção que reduzisse os ciclos entre gerações poderia ser de considerável valor.

Da seleção à colheita de variedades melhoradas, através da produção de sementes em pomares, pode-se levar mais de meio século, mesmo para as espécies de crescimento rápido. Mapas genéticos saturados podem ser construídos com o auxílio de marcadores moleculares como RAPDs, RFLPs, AFLPs ou microssatélites. Nesse caso, as análises de algumas características podem ser realizadas em plântulas jovens – embriões em desenvolvimento, logo após a germinação da semente –, de procedências apropriadas.

Conhecimento & sustentabilidade

Tradicionalmente, uma vez identificados QTLs desejáveis, pode ser necessário um longo tempo de melhoramento intensivo para a restauração das características comerciais, juntamente com o QTL introgredido.

A seleção assistida por marcadores moleculares relacionados a características específicas reduz substancialmente o tempo requerido para o melhoramento. Essa seleção se mostraria especialmente útil, caso esses marcadores pudessem ser diretamente responsáveis pela característica em questão. Isso pode ser obtido usando-se o conhecimento disponível sobre as funções dos genes para sistemas de plantas modelo, através da genômica comparativa.

O conhecimento da base genética da variação inter e intra-específica de características determinadas poderia ser útil no gerenciamento de populações naturais, bem como na exploração sustentável de florestas comerciais.

Boa nova botânica

A compreensão da base genética da diversidade vegetal pode envolver a identificação de diferenças relevantes no controle da expressão ou na função de um grupo restrito de genes, independente da(s) espécie(s) das quais ele(s) se origina(m). O advento da genômica, e mais recentemente da genômica comparativa em plantas, pode tornar possível o estudo integrado de características agriculturalmente desejáveis em plantas anuais e em árvores.

Disponível esse conhecimento, os genes que controlam características de interesse industrial para a silvicultura podem ser identificados e clonados. O gerenciamento e a preservação da variabilidade de ecossistemas florestais naturais também podem ser auxiliados pelas técnicas moleculares de análise genômica comparativa. A associação de técnicas genômicas com a transformação genética de plantas pode permitir a obtenção de florestas melhoradas, com características precisa e pontualmente modificadas.

Marcelo Carnier Dornelas é do Laboratório de Genética Molecular e Melhoramento Florestal, do Departamento de Ciências Florestais da Escola Superior de Agricultura Luiz de Queiroz, USP.

E-mail: mcdornel@carpa.ciagri.usp.br