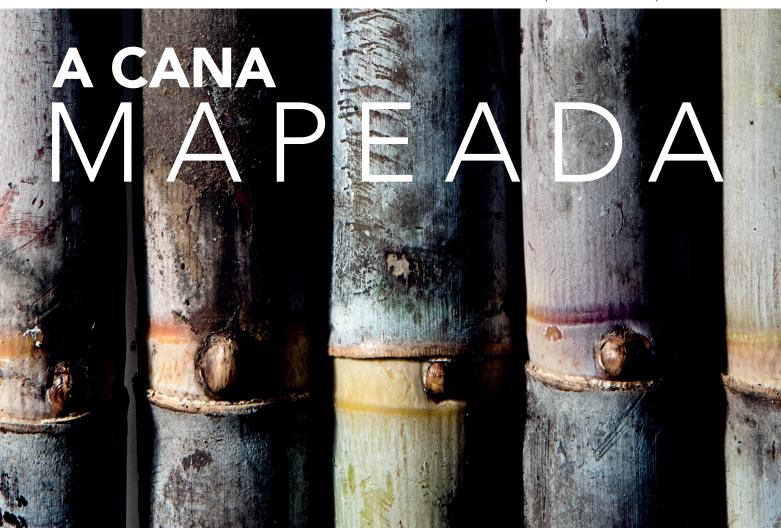
Equipe coordenada por brasileiros encontra mais de 370 mil genes em sequenciamento de cultivar comercial da planta

Eduardo Geraque

m trabalho que dura mais de uma década deu uma contribuição importante para o entendimento da genética da cana-de-açúcar. Um grupo internacional coordenado por pesquisadores do Brasil sequenciou parcialmente o genoma da mais importante variedade comercial dessa planta no país, o cultivar SP80-3280, e encontrou 373.869 genes. O número é 14 vezes maior que a quantidade de genes achada em julho de 2018 por um grupo francês que estudou uma variedade plantada em ilhas do oceano Índico e do Caribe e 10 vezes superior ao determinado por uma equipe chinesa, também no ano passado, para a espécie Saccharum spontaneum, a cana selvagem, não domesticada.

No estudo, a ser publicado na revista científica *GigaScience*, também foram determinadas as possíveis regiões reguladoras que controlam o funcionamento dos genes. "Nosso trabalho foi o mais abrangente porque sequenciamos todo o genoma da cana em vez de apenas

DNA da planta tem 10 bilhões de pares de bases



uma parte, como fizeram estudos anteriores", compara a bioquímica Glaucia Souza, do Instituto de Química da Universidade de São Paulo (IQ-USP), uma das líderes da equipe e membro da coordenação do Programa FAPESP de Pesquisa em Bioenergia (Bioen). O artigo servirá de base para estudos que visam aprimorar a produção de biomassa para gerar energia e de alimentos a partir do melhoramento genético da planta.

Os cultivares de cana são denominados tecnicamente híbridos poliploides. Seu material genético provém de mais de uma espécie e apresenta várias cópias de seus 10 cromossomos básicos. Essa peculiaridade leva seu genoma a ostentar cerca de 10 bilhões de pares de bases, a unidade química que compõe o DNA, mais de três vezes o valor encontrado no Homo sapiens. "Os seres humanos têm duas cópias de cada cromossomo, um herdado do pai e outro da mãe. A cana comercial tem normalmente de 6 a 12 cópias de cada cromossomo", explica a bióloga Marie-Anne Van Sluys, do Instituto de Biociências (IB) da USP, outra líder do grupo.

No estudo, também foi determinado que apenas 12,5% do genoma do cultivar SP80-3280 veio da cana selvagem, conhecida por sua robustez natural, e cerca de 85% de Saccharum officinarum, espécie que o homem começou a plantar há alguns milhares de anos. Um pequeno percentual de seu DNA é fruto da recombinação do material genético dos dois progenitores.

Desde 2008, Souza e Van Sluys trabalham na organização do grupo de pesquisa, que envolve colegas dos Estados Unidos, da China e da Coreia do Sul. A Microsoft Research, em Redmond, no estado norte-americano de Washington, também participou dos trabalhos de sequenciamento. Do ponto de vista técnico, um dos grandes avanços do grupo foi desenvolver métodos de leitura e montagem das longas fitas de DNA em que o gigantesco genoma da cana teve de ser fatiado para ser sequenciado.

A equipe da empresa norte-americana criou algoritmos que permitiram realizar essa tarefa complexa e também usou sua estrutura computacional para processar esse trabalho. "Conseguimos ultrapassar vários obstáculos, todos relacionados à manipulação de um volume grande de dados", explica Bob Davidson, especialista em software para a área de genomas da Microsoft que participou do estudo.

Apesar de o estudo ter mapeado todo o genoma do cultivar SP80-3280, apenas 30% das sequências obtidas, cerca de 3 bilhões de pares de bases, foram montadas de acordo com a ordem em que aparecem em seus cromossomos. Essa parcela do material é a mais importante, pois abriga os genes da planta, que fornecem as instruções para a produção de suas proteínas. Montar um terço do genoma pode parecer pouco, dado o gigantismo e a complexidade do material genético da cana, mas é bem mais do que outros grupos de pesquisa fizeram. No sequenciamento realizado pelo Centro de Cooperação Internacional em Pesquisa Agronômica para o Desenvolvimento (Cirad), os franceses trabalharam apenas com uma cópia de um cromossomo indeterminado. Daí terem encontrado somente 25 mil genes.

Além de ter várias cópias de seus cromossomos, o genoma da cana oferece uma dificuldade extra para ser montado: cada cromossomo tem muitos fragmentos móveis de DNA que se repetem em seu interior, os chamados elementos de transposição. "Esses elementos são muito parecidos. Por isso, ainda não conseguimos ter os genes alinhados nos cromossomos", comenta Van Sluys. O grande desafio agora é ordenar todas as sequências identificadas ao longo de cada cromossomo do cultivar nacional.

esmo não sendo a versão final do genoma da cana, o novo sequenciamento deverá ser útil para os estudos de aprimoramento das variedades plantadas no Brasil. Os pesquisadores descobriram, por exemplo, que existem diferenças importantes em parte das sequências reguladoras dos genes da cana. Essas distinções podem fazer com que uma planta se adapte de forma diferente quando exposta a estresses ambientais, como excesso de salinidade, de calor e de seca.

Como os cultivares comerciais da cana distribuem aproximadamente um terco de seu carbono em sacarose, é importante estudar o metabolismo da produção de açúcar e os principais agentes de sua regulação. Um dos resultados apresentados no trabalho publicado foca exatamente na síntese desse tipo de

carboidrato. O cultivar SP80-3280 tem elementos reguladores particulares envolvidos na produção de sacarose que nunca foram encontrados no ancestral S. spontaneum. Os outros dois terços do carbono da cana plantada vão para estruturas como o caule e as paredes celulares em geral. Devido ao alto teor de lignina, molécula que confere rigidez, essas partes duras podem ser queimadas em caldeiras como combustível.

Nesse contexto, as descobertas científicas obtidas indicam que as sequências genéticas do cultivar brasileiro que regulam o processo de partição do carbono estão nas redes de genes definidas durante o crescimento e a maturação da planta. "Entender todos esses processos que envolvem carbono, açúcar e fibras é fundamental quando se pensa em melhoramento genético", comenta Souza.

Segundo os autores do trabalho, o rendimento mundial da cana, de 84 toneladas por hectare, representa apenas 20% do potencial da planta, estimado em 381 toneladas por hectare. Essa projeção estimula uma corrida internacional para o desenvolvimento tanto de estratégias convencionais de melhoramento da cana quanto por meio da biotecnologia. Com a abordagem mais tradicional, o ganho de produtividade tem sido pequeno, de 1% a 1,5% ao ano.

Esse avanço modesto impulsiona a procura por novas tecnologias. Nesse contexto, o trabalho genômico apresentado agora significa um passo à frente nessa disputa, que se desdobra em questões comerciais. "Além dos franceses e dos chineses, os Estados Unidos também estão agora buscando montar um genoma completo da cana, mas ele ainda não foi publicado", afirma Souza.

Projetos

- 1. Sugarcane signaling and regulatory networks (nº 08/52146-0); Modalidade Projeto Temático; Programa Pesquisa em Bionergia (Bioen); Acordo Fapemig; Pesquisadora responsável Glaucia Souza (USP); Investimento R\$ 4.318.073.60
- 2. Sugarcane genome sequence: Plant transposable elements are active contributors to gene structure variation, regulation and function (nº 08/52074-0); Modalidade Projeto Temático; Programa Bioen; Pesquisadora responsável Marie-Anne Van Sluys (USP); Investimento R\$ 4.190.155,40

Artigo científico

SOUZA, G. M. & VAN SLUYS, M.A. et al. Assembly of the 373K gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world's leading biomass crop. GigaScience. No prelo.