

PELA SAÚDE DAS PESSOAS E DAS PLANTAS

Biólogos, matemáticos e bioquímicos unem conhecimentos para mapear DNA, estudar proteínas e impulsionar a medicina de precisão

RICARDO AGUIAR

A eficácia de um medicamento depende também do organismo de quem o toma. O poder de ação de um fármaco é influenciado pelas características genéticas de cada indivíduo, razão pela qual um mesmo composto pode, às vezes, causar efeitos distintos, e até contrários, em pessoas diferentes. Um exemplo da ação heterogênea de medicamentos foi observado recentemente em um estudo conduzido pelo biólogo Daniel Martins-de-Souza, pesquisador do Instituto de Biologia (IB) da Unicamp. Ele e seus colaboradores na Alemanha avaliaram o desempenho de três medicamentos antipsicóticos em 58 pessoas com esquizofrenia, doença neuropsiquiátrica que atinge 1% da população e é marcada pela ocorrência de alucinações e delírios. Tratados com doses diárias de olanzapina, quetiapina

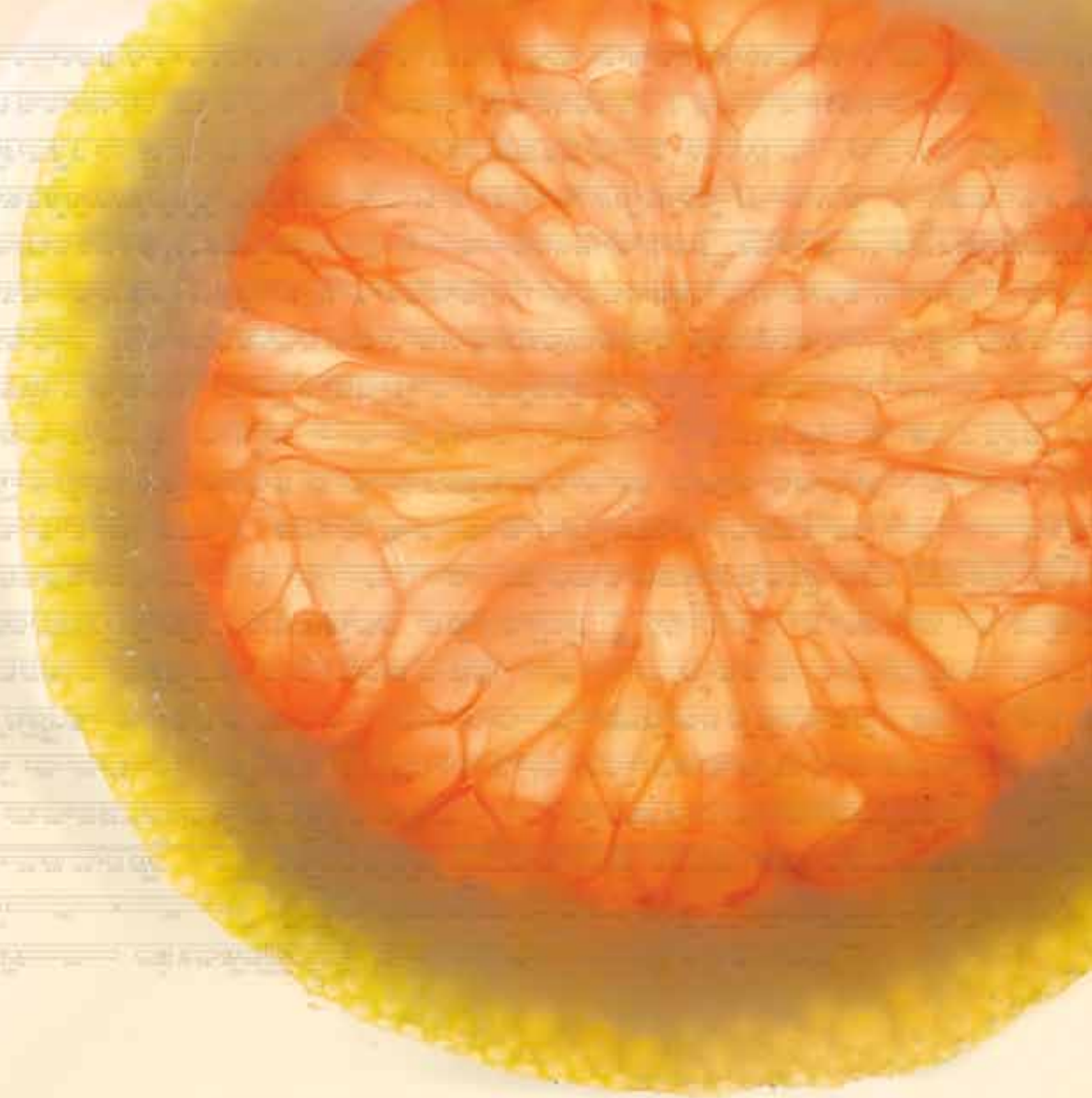
ou risperidona por seis semanas, 36 participantes (62% do total) apresentaram uma redução importante dos sintomas, enquanto os demais não melhoraram.

Os pesquisadores também analisaram o sangue de quem participou do estudo em busca de um possível marcador biológico de resposta ao tratamento, algo capaz de prever se a medicação produziria ou não o efeito desejado. Eles mediram o nível de 985 proteínas e observaram que o grupo que respondeu bem à medicação produzia duas classes de proteínas em concentrações diferentes das pessoas para quem o tratamento não havia surtido o efeito desejado. Os indivíduos que haviam se beneficiado do uso de antipsicóticos produziam níveis mais baixos de proteínas associadas ao metabolismo de ácidos nucleicos e de proteínas que processam outras proteínas, enquanto quem

não respondia bem ao tratamento apresentava concentrações sanguíneas mais elevadas desses dois grupos de proteínas. Combinando o que encontraram em um caso e no outro, Martins-de-Souza e seus colegas chegaram a um conjunto de 42 proteínas que talvez possa indicar quem responderia melhor à medicação, antes mesmo do início do tratamento.

A esperança dos pesquisadores é chegar a um exame de sangue que permita ao médico identificar a medicação mais adequada para cada pessoa de modo a melhorar o índice de sucesso dos tratamentos. “Hoje não há ferramentas que auxiliem o médico a encontrar a medicação mais eficaz”, afirma Martins-de-Souza. “Ele descobre o que funciona na base da tentativa e do erro.”

Ajudar na busca de um marcador biológico de resposta ao tratamento é apenas



O primeiro genoma: de 1997 a 2000, equipes do IB ajudaram a sequenciar a praga dos laranjais

uma das possíveis aplicações da proteômica, área da biologia que estuda o perfil de produção de proteínas dos organismos. “Estudamos o que acontece na esquizofrenia, mas, em princípio, essa estratégia poderia ser utilizada para avaliar o melhor tratamento para qualquer doença.”

Os trabalhos em proteômica ainda são recentes no Brasil, onde a área foi introduzida há pouco mais de uma década, logo após os primeiros estudos de genômica. O IB foi um dos centros nacionais pioneiros nas duas áreas. Antes dos trabalhos em proteômica, seus pesquisadores participaram do Projeto Genoma-FAPESP, que de 1997 a 2000 sequenciou o material genético da bactéria *Xylella fastidiosa*, causadora do amarelinho, ou clorose variegada dos citros (CVC), praga que afeta laranjais do estado de São Paulo e causa grandes prejuízos econômicos.

Esse foi o primeiro genoma de um microrganismo causador de doença em plantas (fitopatógeno) sequenciado no mundo. No projeto, os pesquisadores determinaram a ordem dos 2,7 milhões de pares de bases nitrogenadas (adenina, timina, citosina e guanina, ou A, T, C e G) que compõem o material genético (DNA) da *Xylella*. “Esse projeto foi um divisor de águas, pois equiparou as pesquisas brasileiras feitas nessa área às pesquisas internacionais”, diz Martins-de-Souza.

“Na época, o Brasil tinha poucas iniciativas de sequenciamento de genes, nenhuma tão grande como essa”, conta o biólogo Paulo Arruda, professor do IB e um dos coordenadores do projeto. Cerca de 190 pesquisadores de 35 laboratórios do estado de São Paulo participaram do projeto, no qual interagiam por meio de uma plataforma chamada Organização

para Sequenciamento e Análise de Nucleotídeos (Onsa, na sigla em inglês). No projeto, equipes de diferentes instituições aprenderam técnicas de sequenciamento, o que acelerou significativamente o desenvolvimento da genômica no país. “Realizar o projeto nos permitiu acompanhar a evolução da área de forma extremamente efetiva”, diz Arruda, que hoje usa a genômica para tentar desenvolver plantas resistentes à seca, como milho e soja, em parceria com a Empresa Brasileira de Pesquisa Agropecuária (Embrapa).

BIOINFORMÁTICA

À medida que se acumulava a enorme quantidade de informações do genoma, foi preciso desenvolver a capacidade de analisar e processar os dados, tarefa da bioinformática, área que usa estratégias computacionais para analisar dados bio-

lógicos. Na época do Genoma-FAPESP, a bioinformática ainda era uma área muito recente no Brasil e chegou-se a pensar em trazer especialistas do exterior para colaborar com o projeto. Mas dois jovens pesquisadores do Instituto de Computação (IC) da Unicamp à época, João Meidanis e João Carlos Setubal, trabalhavam com isso e chamaram a atenção dos coordenadores do Genoma-FAPESP.

Com a entrada de Meidanis e Setubal no projeto, o IC começou a desenvolver expertise em uma nova área, assim como já havia feito quase três décadas antes, ao formar a primeira turma de cientistas da computação no Brasil. A Unicamp foi a segunda universidade a criar um curso de ciências da computação, em 1969, um mês depois da Universidade Federal da Bahia.

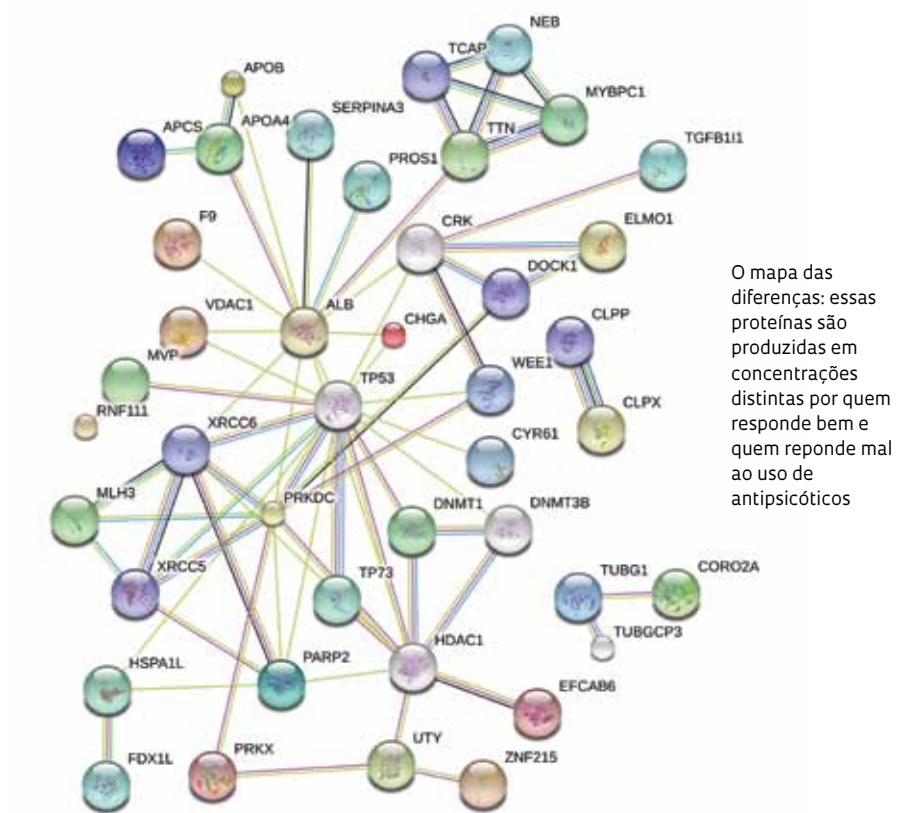
Os dois pesquisadores implementaram técnicas e desenvolveram softwares que permitiram o reagrupamento de trechos do DNA da bactéria sequenciados por laboratórios diferentes, como peças de um quebra-cabeça gigante. “Além disso, criamos sistemas de transmissão de dados via web para que todos os pesquisadores pudessem acompanhar o andamento do projeto a partir de seus próprios laboratórios”, conta Meidanis. “No Brasil, nada semelhante havia sido feito antes.”

Em 2000, os bioinformatas receberam com outros pesquisadores a Medalha do Mérito Científico e Tecnológico do Governo do Estado de São Paulo pela sua participação no Genoma-FAPESP. Em 2009, também foram homenageados com o prêmio Distinguished Innovators Award da Business Software Alliance pelo trabalho com bioinformática no mesmo projeto.

PROTEOMA

Concluído o sequenciamento da *Xylella*, o passo seguinte foi a realização do Projeto Genoma Funcional da bactéria. Vinte e um laboratórios começaram a investigar a função dos quase 2,8 mil genes e das proteínas por eles produzidas com o objetivo de compreender como a *Xylella* infecta as laranjeiras e cresce no interior dos vasos que permitem o transporte de nutrientes.

Um desses grupos foi coordenado pelo biólogo José Camillo Novello, do IB. Sua equipe mapeou o proteoma total da bactéria e descobriu que a *Xylella* não dispõe de um mecanismo que permi-



O mapa das diferenças: essas proteínas são produzidas em concentrações distintas por quem responde bem e quem responde mal ao uso de antipsicóticos

te produzir proteínas de forma rápida, comum em outras bactérias. Essa característica, identificada pelo biólogo Marcus Smolka, atualmente professor na Universidade Cornell, nos Estados Unidos, ajudou a explicar o crescimento lento do microrganismo, que demora de oito a 10 horas para se dividir em dois. Só para comparação, a *Escherichia coli* não leva mais de 20 minutos para se duplicar. O grupo de Novello identificou 142 proteínas e estudou em detalhe 30 delas, potenciais alvos para o desenvolvimento de compostos contra o amarelinho.

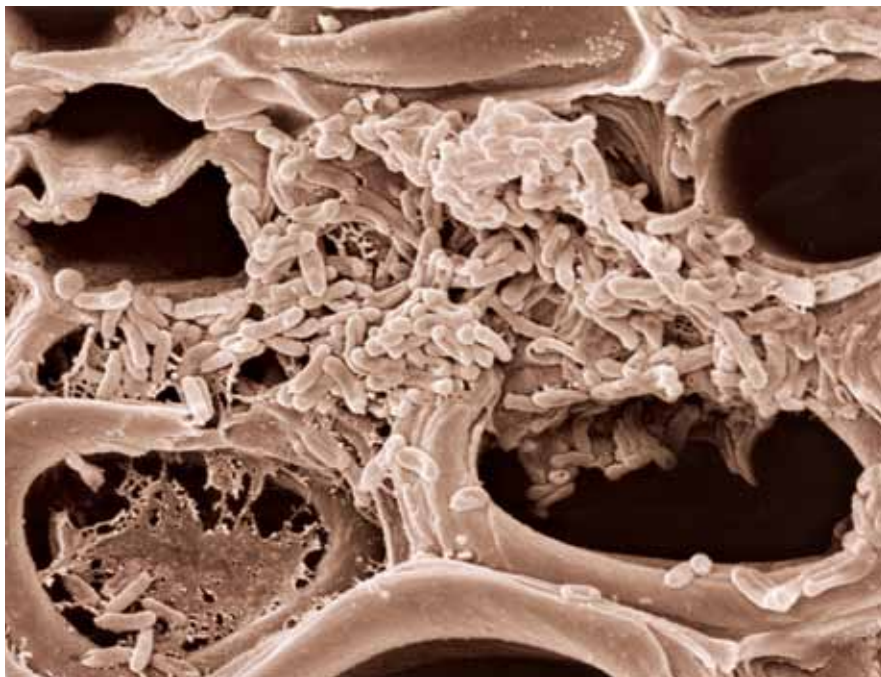
O conhecimento gerado nesses projetos propiciou a produção de testes para detectar a presença da bactéria em mudas de plantas. Esses testes foram importantes para o controle da doença e para ajudar a reduzir as perdas – muitas vezes o amarelinho só se manifesta anos depois que a muda é plantada.

Iniciado em 2000 e concluído com sucesso dois anos depois, o Genoma Funcional da *Xylella* foi o primeiro projeto a gerar um proteoma completo na América Latina. “Ele abriu as portas para novas pesquisas em proteômica no Brasil”, diz Novello.

O conhecimento conquistado nos anos 1990 com o sequenciamento do material genético de plantas, animais

e microrganismos auxiliou na investigação do genoma humano. Vencidas as barreiras tecnológicas iniciais, hoje a genômica é usada corriqueiramente na área médica. Em 2015, cinco Centros de Pesquisa, Inovação e Difusão (Cepid) financiados pela FAPESP se uniram e criaram a Brazilian Initiative on Precision Medicine (BIPMed), a primeira base de dados genômicos do país, que integra um consórcio internacional de bancos genômicos associados à saúde humana. O objetivo é desenvolver terapias tailhadas para cada paciente.

Estão disponíveis na base da BIPMed o exoma completo de 29 pessoas. O exoma é a parte do genoma que codifica proteínas e, portanto, qualquer mutação que venha a sofrer pode ter como consequência o aparecimento de uma doença. Os grupos que integram a BIPMed já identificaram 209 variantes genéticas que não haviam sido descritas, mas ainda não se sabe se essas variações genômicas estão relacionadas a doenças. A base também abriga informações sobre mais de 1 milhão de marcadores moleculares de polimorfismo de base única de 200 indivíduos de uma população de referência. Essas variações não costumam causar doenças, mas podem fornecer



Colônia de bactérias *Xylella fastidiosa* obstrui vasos condutores de seiva em laranja

informações sobre a estrutura genética da população. “Essa iniciativa dará maior visibilidade aos dados brasileiros na comunidade científica internacional”, conta o físico Munir Skaf, pesquisador do Instituto de Química (IQ) da Unicamp e coordenador do Center for Computational Engineering and Sciences (CCES), um dos Cepids envolvidos na BIPMed.

CRIAÇÃO DE EMPRESAS

Um dos desdobramentos do Genoma-FAPESP foi a criação de empresas a partir da universidade. Observando o potencial de aplicação na indústria das pesquisas em genômica e proteômica, Arruda e Setubal se uniram a três outros pesquisadores – Ana Cláudia Raseira da Silva, João Paulo Kitajima e Jesus Ferro – e fundaram em 2002 a Alellyx, empresa de pesquisa e desenvolvimento de produtos de biotecnologia e geração e comercialização de patentes na área de genômica aplicada. Por meio da manipulação de genes, a empresa conseguiu conferir a uma variedade de cana-de-açúcar resistência a uma doença conhecida como mosaico. Em 2008, a Alellyx foi comprada pela multinacional Monsanto.

Pouco mais de um mês depois da criação da Alellyx, Meidanis e quatro alu-

nos – Zaroni Dias, Guilherme Pimentel Telles, Marília Dias Vieira Braga e Alexandre Corrêa Barbosa – fundaram a empresa Scylla para atuar na área de bioinformática. Entre outras realizações, a empresa desenvolveu o Sinbiota 2.1, a segunda versão do software usado pelo programa Biota-FAPESP. A atualização lançada em 2013 integrou a ferramenta Google Maps ao programa e permite aos pesquisadores registrar de maneira simples a localização de plantas e animais do estado de São Paulo. Até o momento, 404 usuários de 154 projetos já registraram mais de 126 mil espécimes no software.

Atualmente, a principal demanda da Scylla é em metagenômica, área da ciência que realiza o sequenciamento de genomas encontrados em determinado ambiente. O desafio, no caso, é identificar os diferentes organismos que vivem ali.

“Também temos interesse em medicina personalizada, área promissora que pode, no futuro, se tornar a maior demanda por parte de empresas”, diz Arruda. Para ele, a genômica e a proteômica se tornaram essenciais para responder a perguntas biológicas da maneira mais precisa. “A genômica deixou de ser uma barreira para se transformar em uma ferramenta”, afirma. ■

ARCHC, UM SOFTWARE PARA TESTAR PROCESSADORES

Está cada vez mais complicado produzir os circuitos eletrônicos do processador de um computador. O número de transistores que integram um chip cresceu tanto nos últimos tempos – os atuais podem ter mais de 10 milhões de transistores por milímetro quadrado – e as funções desempenhadas por eles se tornaram tão complexas que é inviável fabricar uma peça primeiro e só testá-la depois. Agir dessa forma pode significar perda de tempo e dinheiro.

Acompanhando essa tendência de miniaturização dos componentes há bastante tempo, os pesquisadores Sandro Rigo, Guido Araújo e Rodolfo Azevedo, todos do IC, decidiram desenvolver no início dos anos 2000 uma linguagem que permite aos projetistas e engenheiros simularem o funcionamento de um processador antes de sua fabricação.

Essa linguagem, a ArchC, é, segundo Rigo, a única criada no Brasil e uma das poucas produzidas no mundo para essa função e no formato *open source*, que permite ao usuário baixar o programa gratuitamente e alterá-lo de acordo com suas necessidades. Desde que foi criada, a ArchC já foi – e continua sendo – usada por pessoas e instituições de cerca de 35 países, incluindo universidades como a de Cambridge, na Inglaterra. “Estimo que a ArchC deva ter milhares de usuários em todo o mundo, o que é muito para um nicho tão específico”, diz Rigo.

Em 2016, a linguagem ArchC foi selecionada para o Google Summer of Code, programa do Google que financia alunos para trabalharem no desenvolvimento de projetos *open-source* selecionados mundo afora. Neste ano o programa recebeu cerca de 30 propostas de estudantes e escolheu as duas melhores para aprimorar a ArchC.