

Bioinformática organiza sequenciamento do exoma

Pesquisador da FCM usa ferramentas computacionais para encontrar alterações que possam ser associadas a doenças

CARMO GALLO NETTO
carmo@reitoria.unicamp.br

Várias doenças podem ser explicadas através da informação contida dentro do genoma de cada indivíduo. Mais especificamente, existe dentro do genoma uma região, que corresponde a menos que 2% de sua totalidade, chamada de exoma, tida como potencialmente responsável por armazenar as informações que determinam a composição e regulação dos mecanismos celulares de todo o corpo. Murilo Guimarães Borges, graduado em física-médica pela Unicamp, desenvolveu trabalho em nível de mestrado em que aplica métodos computacionais para organização das sequências de DNA que compõem o exoma de forma a localizar alterações nesta região que possam vir a ser associadas aos mais variados tipos de doenças.

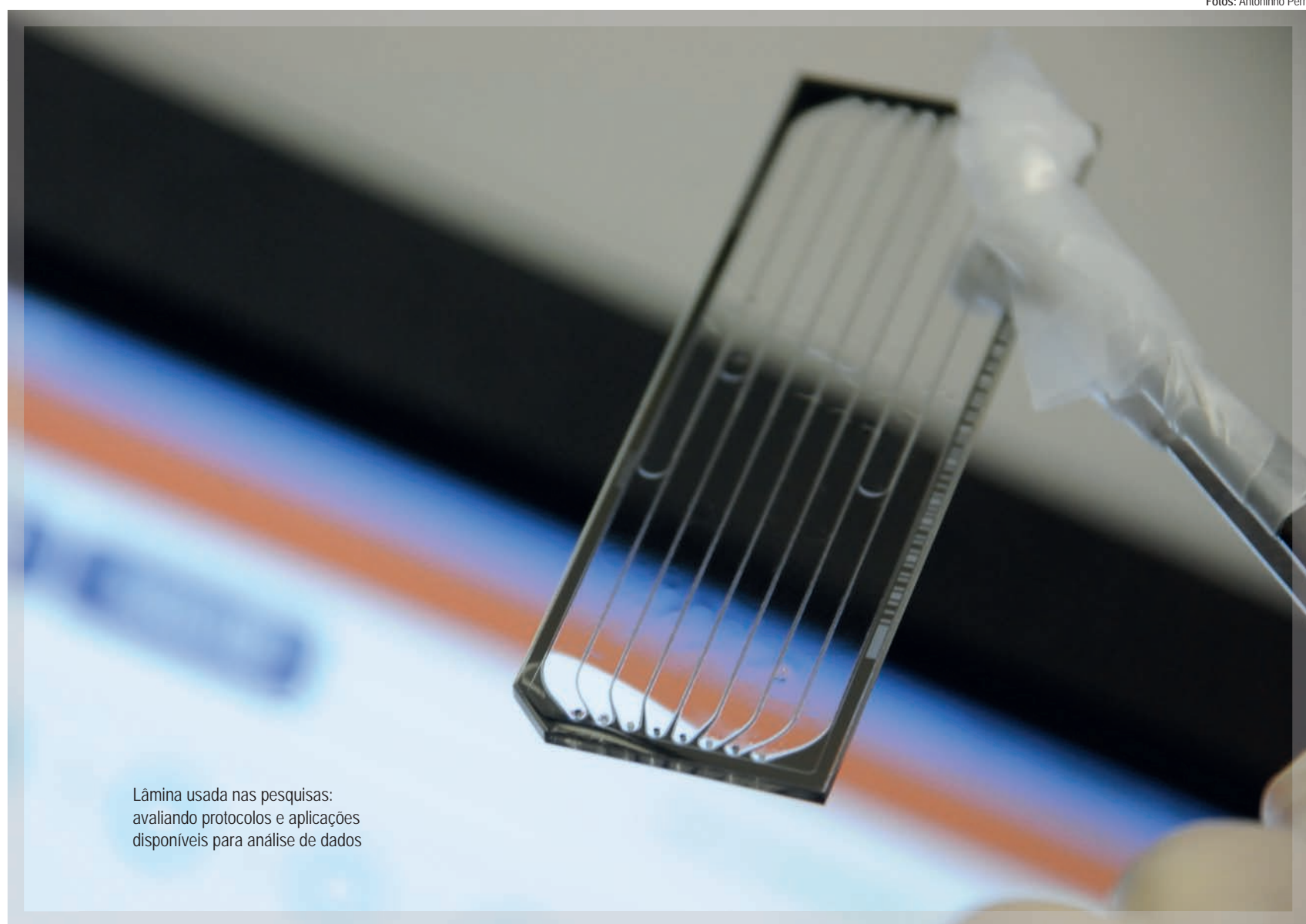
Borges explica que a organização dessas sequências se assemelha à montagem de um quebra-cabeças em que se pode permitir que algumas peças fiquem fora do lugar. São elas que tornam os indivíduos únicos. O desafio é verificar se essas peças deslocadas correspondem a uma variação normal e aceitável em relação àquelas que são patológicas. Não é uma constatação simples. No trabalho, o pesquisador desenvolve formas mais eficientes de encontrar as respostas.

O DNA – ácido desoxirribonucleico –, do ponto de vista químico, é um imenso polímero constituído de unidades mais simples, as bases nitrogenadas. Ele contém as instruções genéticas que coordenam o desenvolvimento e funcionamento dos indivíduos e transmitem suas características. Seu principal papel é armazenar as informações necessárias para a construção de proteínas.

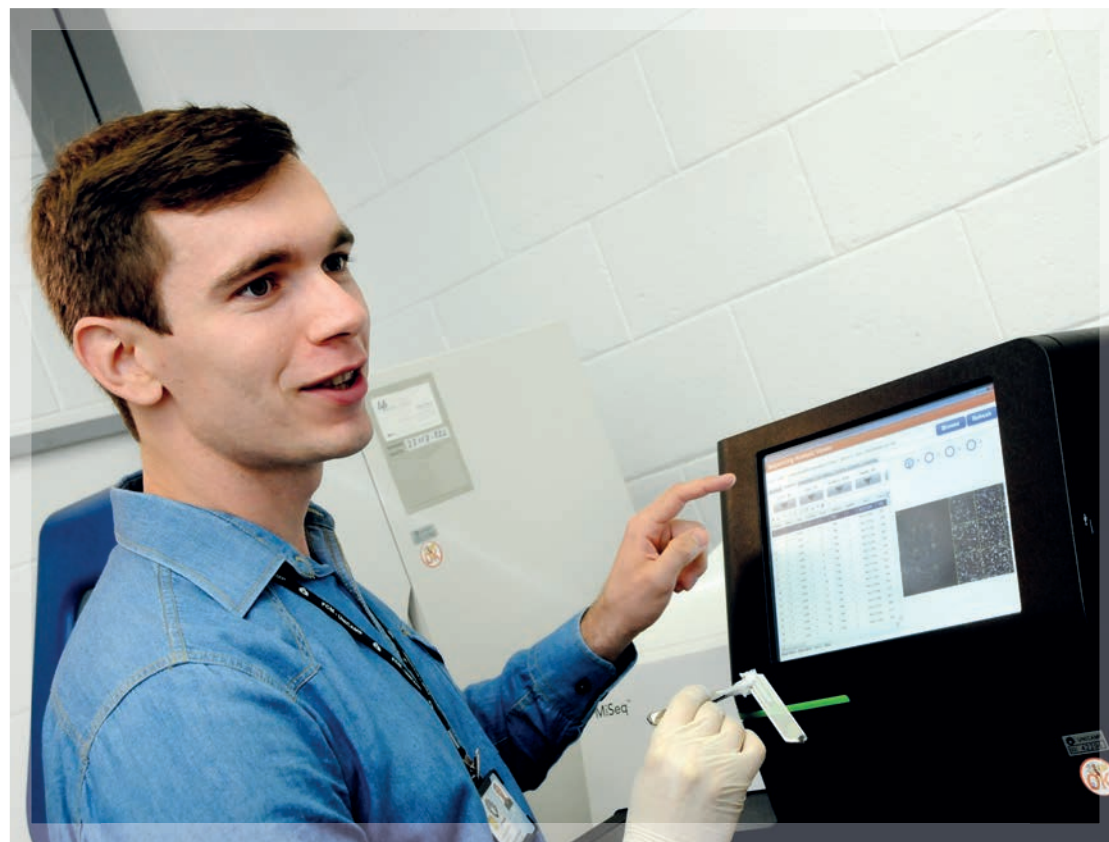
Ele é constituído por duas fitas complementares, formadas por moléculas de adenina, citosina, guanina e timina. Cada tipo de base de uma cadeia se liga apenas a uma base complementar da outra cadeia. A informação contida nesta molécula de DNA pode conter alterações que resultarão na formação de proteínas que não funcionam como deveriam. Algumas destas variações são responsáveis por doenças.

O TRABALHO

O DNA humano é constituído por mais de três trilhões de bases. Em uma reação de sequenciamento, em que se utilizam aparelhos para sua leitura, não se consegue ler todo o genoma de uma só vez. Então se utiliza um sequenciador que lê porções menores da informação que o constituem. Isso é basicamente o que acontece numa reação de ressequenciamento. Finalmente, procuram-



Lâmina usada nas pesquisas: avaliando protocolos e aplicações disponíveis para análise de dados



Murilo Guimarães Borges, autor da dissertação: organização de sequências se assemelha à montagem de um quebra-cabeça

-se as diferenças com o genoma de referência porque elas, que individualizam os indivíduos, podem estar também relacionadas a doenças.

No trabalho o autor faz a aplicação de protocolos e métodos em bioinformática para a análise de sequenciamento de exomas humanos. Na bioinformática se usam computadores e softwares para coletar, analisar e interpretar informações biológicas. A análise através dela é essencial para lidar com o alto volume de dados gerados e realizar a ligação entre eles e o experimento biológico.

Os objetivos do estudo foram avaliar protocolos e aplicações disponíveis que pudessem ser utilizadas na análise dos dados gerados pelo sequenciamento de exomas humanos; aplicar e aperfeiçoar protocolos e aplicações disponíveis para prever variantes potencialmente patológicas a partir de dados gerados pelo sequenciamento de exomas humanos; caracterizar o impacto das diferenças técnicas e populacionais na cobertura resultante da captura do exoma.

A propósito, Murilo esclarece: “Procurei encontrar uma forma, um método que, com base nas informações fornecidas pelo se-

quenciador, possibilitasse chegar a uma lista das variantes encontradas. Não nos cabe diagnosticar a doença, mas apontar diferenças no exoma e, a partir daí, possibilitar que um especialista tente localizar que alteração é potencialmente responsável por ela”.

O desafio é montar o quebra cabeça para identificar as diferenças que podem ou não ser patológicas de modo a garantir que as alterações observadas de fato ocorreram no genoma analisado e não decorrem do emprego do computador ou do trabalho de bancada no laboratório. “Ao fornecer as alterações no DNA há necessidade de garantir ao máximo que elas sejam reais e não resultem da técnica utilizada, e isso é o que tentamos fazer em nossas análises em bioinformática”, esclarece o autor.

RESULTADOS

Realizado o sequenciamento é determinante um contraste com a literatura, onde se encontram informações de que malefícios essas pequenas diferenças podem vir a provocar no indivíduo, em que porcentagens podem ser danosas, ou seja, o grau de benignidade ou malignidade das alterações encontradas.

Murilo conclui: “Aplicamos o protocolo nas amostras estudadas e chegamos a resultados que podemos fornecer para os pesquisadores de forma anotada, com informações relevantes para guiar experimentos futuros”. Ele identificou também algumas oportunidades. Assim é que, uma variante em relação ao genoma de referência suscita uma questão: esta diferença não estaria relacionada ao fato de se estar comparando o genoma de um brasileiro com o de um indivíduo de outro background genético? É uma possibilidade que deve ser considerada. Este fato o levou a julgar de extrema importância a montagem de um banco de variantes da população brasileira. Mesmo porque uma mutação frequente na população do Brasil pode não ser patogênica e podem estar sendo feitas associações equivocadas em relação à genética de outros povos. Daí a importância da geração de banco das alterações comuns na população brasileira.

Ele identificou, ainda, que é possível modificar as sondas utilizadas para selecionar o exoma incluindo nelas diferenças relacionadas à população considerada, porque essas sondas podem não ser as mais eficientes na nossa população. Em decorrência, uma empresa que produz kits para captura de exomas poderia se beneficiar dessas informações para construir sondas que contribuíssem para maior eficiência da metodologia.

O protocolo apresentado mostra a sequência de passos e de programas computacionais utilizados para, a partir das sequências brutas que saem do sequenciador, chegar às listas de variantes anotadas fornecidas aos pesquisadores.

Resumindo, diz Murilo: “Partindo do DNA, selecionamos o exoma e a partir dele, que constitui 2% das sequências do DNA, utilizamos vários passos que permitiram, com base nas sequências brutas, chegar às variantes no indivíduo analisado em relação ao genoma adotado como referência. Tentamos mostrar que essas variantes são verdadeiras e que não resultavam de um viés da técnica utilizada. Adicionamos, ainda, várias informações a essas variantes, que constituem as anotações. Então, a partir das sequências brutas, chega-se a uma lista de variantes anotadas para que o pesquisador, responsável pelas sequências, possa tentar associar cada uma delas com a doença ou condição estudada”.

Publicação

Dissertação: “Aplicação de protocolos e métodos em bioinformática para análise de sequenciamento de exomas humanos”

Autor: Murilo Guimaraes Borges

Orientadora: Iscia Teresinha Lopes Cendes

Unidade: Faculdade de Ciências Médicas (FCM)