

Parceiros e Colaboradores

Buscar

DIKITE SUA PESQUISA

MELHORAMENTO E DIVERSIDADE GENÉTICA

Por que sequenciamos genomas na Embrapa Agroenergia?

sexta-feira, dezembro 15, 2017



Muito se fala sobre genômica e genoma humano, milhares ou milhões de reais são investidos anualmente nesta área pelo mundo, e várias revistas científicas trazem novos genomas em suas capas. Mas, o que é genoma? Por que é importante entendê-los? E, por onde começar?

Genoma é o conjunto do material genético e hereditário dos organismos vivos, codificado em seu DNA (em alguns casos, no RNA). A Genômica é a ciência que estuda os genomas. Nosso genoma, por exemplo, é muito complexo, e organizado em cromossomos. Cada um dos cromossomos é

composto por uma longa cadeia de DNA (dupla hélice ou dupla fita), e o DNA é a forma como quase todos os seres vivos do nosso planeta armazenam e passam adiante as informações genéticas (em alguns casos é o RNA, como em vírus). No DNA estão os genes, utilizados como “moldes” para a produção de proteínas, e também outras estruturas, incluindo um complexo sistema de regulação gênica.

Resumindo a história, a partir da sequência de um genoma, diversas análises podem ser realizadas, como a anotação das estruturas, a comparação entre diferentes genomas, as análises filogenéticas, de vias metabólicas, de regiões reguladoras e muitas outras. De modo geral, o genoma é a base necessária para uma série de estudos posteriores.

Desde que o sequenciamento de genomas maiores, como o humano, ficou economicamente viável, a obtenção do draft do genoma tornou-se um dos primeiros passos na pesquisa de organismos relevantes para determinadas linhas de pesquisa. Draft (rascunho) é uma versão incompleta da sequência de um genoma, e é a versão mais comum, pois obter um genoma absolutamente completo ainda é difícil e caro, e só se justifica para genomas mais simples e/ou para organismos de alto valor (como organismos modelo).

Mas como se chega ao draft de um genoma? A obtenção de um novo genoma começa pelo seu sequenciamento, processo de identificação da sequência de nucleotídeos do DNA (Adenina, Guanina, Citosina e Timina), que inclui principalmente: escolha do material; preparação do DNA; sequenciamento; controle de qualidade das sequências; montagem; e anotação das estruturas.

O primeiro passo é a escolha do material biológico adequado, como uma planta importante num programa de melhoramento genético, ou uma cepa de levedura com capacidade diferenciada de produção. De modo geral, a escolha está relacionada com um indivíduo diferenciado de uma espécie importante para uma linha de pesquisa, o que indica normalmente relação com produtividade agrícola, industrial ou saúde. Por exemplo, na Embrapa Agroenergia estamos trabalhando com genomas de plantas relacionadas a bioenergia, como caiaué (dendê americano), macaúba e pinhão-manso, e microrganismos relacionados a processos de produção de energia, como leveduras e microalgas.

Após a coleta do material biológico, é realizada a extração do DNA, que é preparado, avaliado e enviado para o sequenciamento. Existem diferentes tecnologias de sequenciamento. Todas geram cópias de pedaços do DNA original, sendo necessária posterior montagem dos fragmentos para obtenção da sequência original. A capacidade de gerar sequências (tecnologias e equipamentos)

aumentou consideravelmente com as demandas dos projetos do Genoma Humano, publicados em 2001 nas duas mais importantes revistas científicas da área (Nature e Science). Conforme a quantidade de dados biológicos gerados aumentava, também crescia a importância dos bioinformatas para a análise destas grandes quantidades de informação. Grandes quanto? Estamos falando de vários arquivos com milhões de linhas e vários Giga Bytes (GB) cada um. Por exemplo, um arquivo de 60 GB (que nem abre no Windows) tem cerca de 823 milhões de linhas, correspondendo a 205,8 milhões de sequências, e um bom Servidor Linux demorou nove (9) minutos apenas para contar o seu número de linhas. Grandes assim.

Além da grande quantidade, estes dados não são 'perfeitos', isto é, parte das sequências normalmente apresentam problemas, como: baixa qualidade, resíduos da construção das bibliotecas (parte da preparação do DNA) e do sequenciamento, ou contaminações. Por isto, o controle de qualidade (QC, de Quality Control) dos dados de sequenciamento é essencial.

O primeiro passo é analisar e visualizar a qualidade dos dados brutos. São verificadas informações como a qualidade (na identificação dos nucleotídeos, ou seja, a chance de que esta identificação esteja correta), os resíduos de adaptadores e a proporção entre os nucleotídeos. De acordo com as características de cada conjunto de dados, outras ferramentas são utilizadas para filtrar os dados. Nesta fase, são retiradas, por exemplo, sequências com baixa qualidade, contaminações, e trechos com resíduos de adaptadores ou vetores. Quando necessário, é verificado o pareamento das sequências e o tamanho dos insertos.

O processo de Controle de Qualidade tem muitos detalhes, variando com o tipo de tecnologia, de sequenciamento solicitado, de organismo, de software a ser utilizado na montagem, entre outros. Trata-se de uma tarefa complexa e que demanda pessoal especializado e infraestrutura robusta de hardware e software, pois além dos processos serem demorados, é necessário que sejam executados várias vezes para verificar melhores parâmetros.

Tudo isto para conseguir aproveitamento máximo dos dados, e em 2017 a equipe do LBB (Laboratório de Bioinformática em Bioenergia - <https://lbb.cnpae.embrapa.br>) está trabalhando em sequenciamento e montagem de novo de vários genomas, sendo pelo menos três de plantas, três de algas, e três de leveduras, com previsão de novos genomas para 2017, incluindo o resequenciamento de cerca de 40 genomas de plantas. Montagem? Falaremos mais detalhadamente sobre a montagem de genomas e a anotação de suas estruturas no próximo artigo, não perca. Se for montar seu próprio genoma, não

se esqueça de fazer um bom controle de qualidade!

Embrapa Agroenergia

Mais informações sobre o tema

Serviço de Atendimento ao Cidadão (SAC)

www.embrapa.br/fale-conosco/sac/



MACAÚBA

Portal dedicado à Macaúba. Aqui abordamos tudo sobre essa promissora espécie.

VEJA TAMBÉM:



ESTUDOS GENÉTICOS EM
POPULAÇÕES NAT...

[ARTIGO] NATUREZA DOS
METABÓLITOS P...

MACAUBA: A PROMISING
TROPICAL PALM ...

0 COMENTÁRIOS

Agradecemos seu comentário!

Comentar como:

MAIS RECENTES

MAIS ANTIGOS

MAIS LIDAS DA SEMANA

CATEGORIAS



Os 6 Benefícios da Macaúba Para Saúde



As mil e uma utilidades da Macaúba



Empresa compra óleo de cozinha usado para exportar

Aspectos Gerais (49) Cultivo (76) Endocarpo (8) Farelos (25)
Macaúba (150) Melhoramento e Diversidade Genética
(42) **Mercado (2616)** Óleos (80) Produtos (63)
Vídeo (163)

Created By Sora Templates & Blogger Templates Personalizado e mantido por ENTER.elo