

# Melhoramento genético **HIGH-TECH**

Por: Wagner Arbex - Doutor em Engenharia de Sistemas e Computação, Engenharia Brasileira de Pesquisa Agropecuária

AD

Há pouco tempo, ninguém seria capaz de imaginar que, em um programa de auditório em televisão aberta, um participante seria desafiado mediante procedimentos científicos e de laboratório. Hoje, contudo, não é incomum que apresentadores de sucesso "organizem o circo" no qual, juntamente com a plateia aos brados, desafiam um possível progenitor, que parece estar entregue aos leões – ou seria aos ratinhos: "faz DNA?". Sem dúvida, isto é pesquisa científica aplicada como prática, e não se vê.

**E**ssa observação foi feita durante um encontro acadêmico sobre genômica e melhoramento genético na Universidade Federal de Viçosa, há quase oito anos, quando o palestrante, que também é o autor desse texto, estava intencionado a mostrar onde a computação e a genômica haviam chegado, naquele momento.

A proposta de um exame de DNA para verificação de paternidade é um exemplo claro da união de técnicas da genômica, com recursos da química e da bioquímica, entre outras, em um procedimento operacionalizado por ferramentas surgidas a partir do avanço tecnológico, a partir de fundamentos da computação e seus recursos.

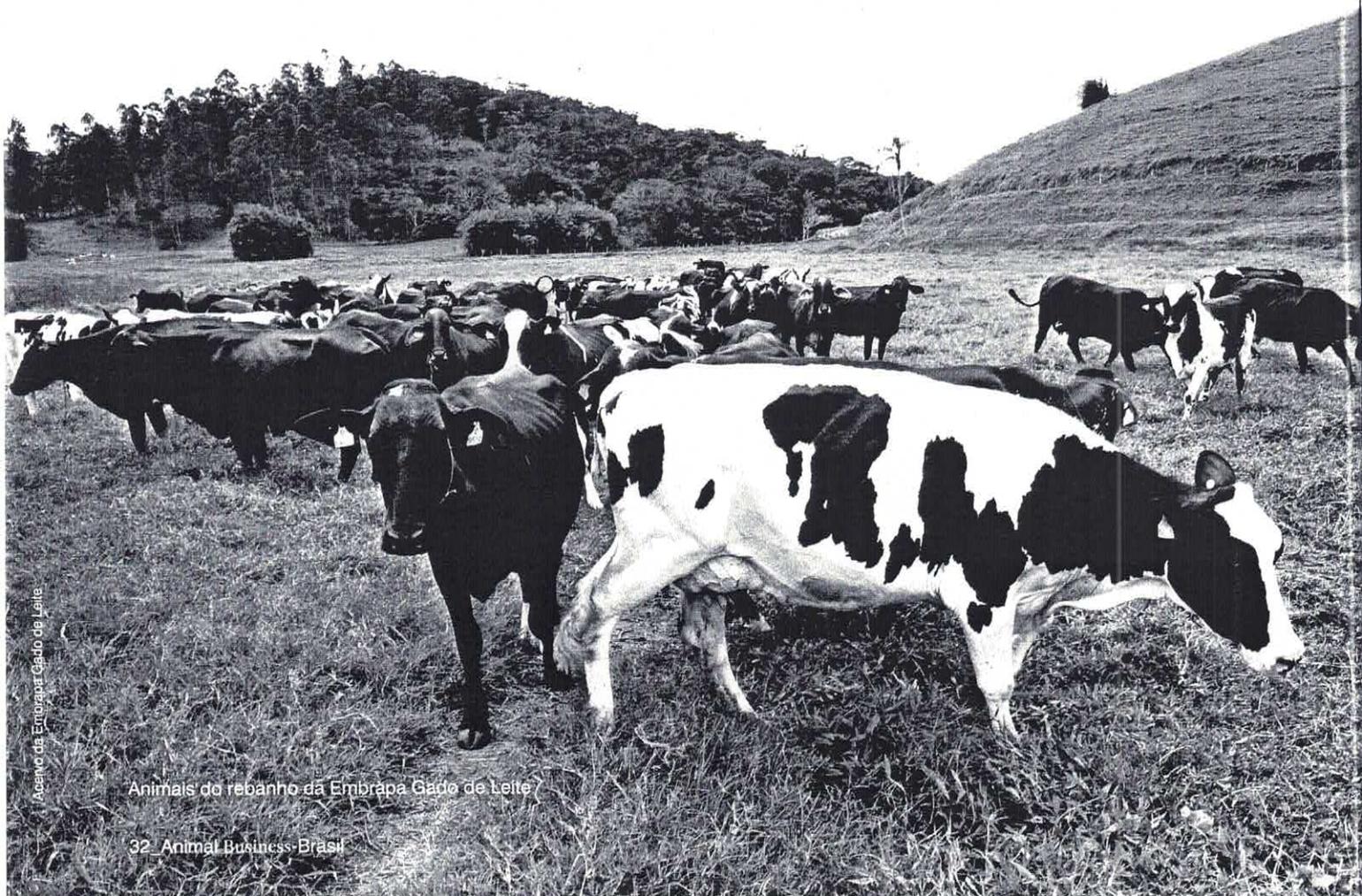
Na época, a união da genômica e da computação não era desconhecida, pois já existiam diversos trabalhos desenvolvidos pela comunidade acadêmica onde, essas duas áreas, caso não estivessem juntas, não seria possível o desenvolvimento de tais trabalhos.

À primeira vista, a junção de áreas tão diferentes causava alguma estranheza no público presente e, apesar do imenso avanço e das inúmeras aplicações dessas áreas, essa mesma percepção ainda é notada até hoje.

Entretanto, atualmente, a maioria dos profissionais que atuam na genômica e, ainda, outra grande parcela dos profissionais da computação, percebem claramente que suas atividades formam uma espécie de ciclo, no sentido de que uma é a continuidade da outra.

Dessa forma, alguns métodos utilizados na ciência da computação tornaram-se essenciais para a solução de problemas da genômica. Mas, em contrapartida, somente a aplicação desses métodos, para solucionar problemas dessa natureza, possibilita o desenvolvimento de novos métodos e recursos da própria ciência da computação.

As ações da pesquisa em genética e genômica, que antes se encontravam quase totalmente limitadas “às bancadas”, já não existem mais sem



que sejam complementadas com procedimentos computacionais. Assim, os dados obtidos nas bancadas podem ser devidamente identificados, armazenados, organizados, agrupados e classificados para que, posteriormente, possam ser recuperados, apresentados e, ainda, utilizados como um novo atributo de informação.

Dessa forma, o que antes podia parecer estranho, pela distância que existia entre os dois conceitos, hoje, pode ainda não ser totalmente entendido, mas pela proximidade e "mistura" das aplicações desses conceitos.

## Genética e computação

Historicamente, os trabalhos e avanços do melhoramento genético sempre estiveram relacionados à computação – ou melhor, à disponibilidade dos recursos adequados para as atividades de melhoramento genético – pois, devido à natureza dos trabalhos de pesquisa, quer seja por meio de cruzamento ou de seleção, os conjuntos de dados envolvidos nas ações de pesquisa e desenvolvimento importam tanto na quantidade quanto na qualidade dos mesmos.

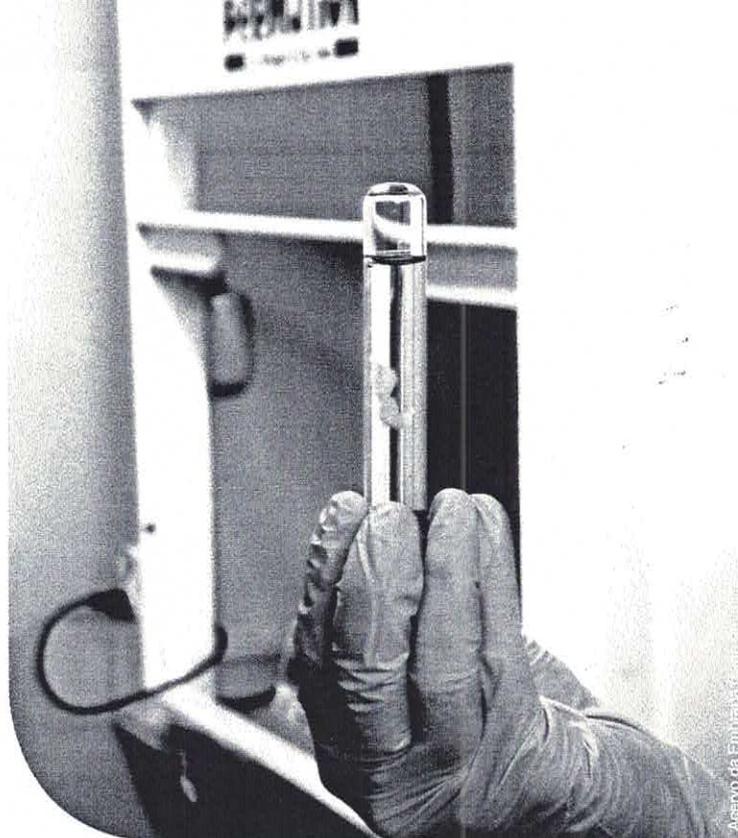
Por exemplo, atividades em prol do melhoramento genético animal, quase sempre, devem envolver a realização de avaliações tradicionais dos animais, conhecidas como avaliações genéticas para a identificação e seleção de animais com potencial genético diferenciado e próprio para o fenótipo de interesse.

Por sua vez, essas avaliações de animais, que são realizadas com o uso de dados fenotípicos e de pedigree dos animais, podem – aliás, podem e devem – envolver um número considerável de informações pois, a confiabilidade dos resultados dessas avaliações baseiam-se fortemente na quantidade de informações obtidas de cada animal e de seus descendentes e antecedentes.

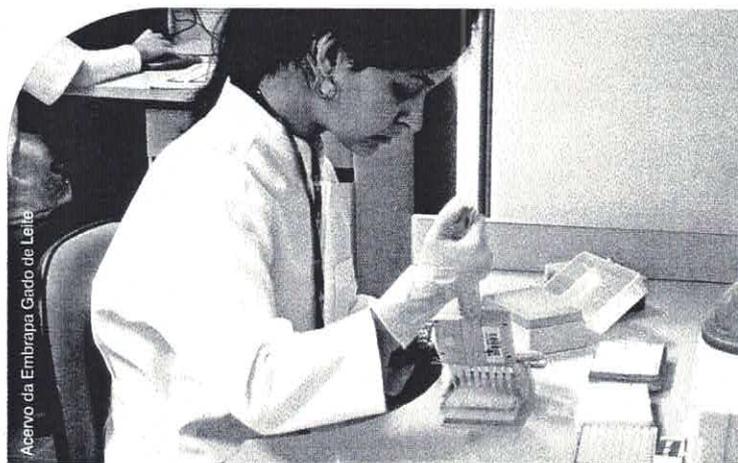
Para essas avaliações genéticas, a computação importa não só pelos métodos de tratamento dos dados, mas também pelas estratégias para execução das avaliações em si.

Os modelos matemáticos para a execução das avaliações genéticas, atualmente utilizados em programas de melhoramento por todo mundo, só foram desenvolvidos graças aos modelos e estratégias computacionais para suportá-los.

Isto é, mais uma vez, os conceitos e as técnicas próprias das áreas diretamente relacionadas



DNA "precipitado" durante o processo de extração



Preparo de amostra no Laboratório de Genética Molecular

Historicamente, os trabalhos e avanços do melhoramento genético sempre estiveram relacionados à computação



Foto: A. E. Moraes/Embrapa Gado de Leite

Animais do rebanho da Embrapa Gado de Leite no Complexo Multiusuário de Bioeficiência e Sustentabilidade da Pecuária

ao objetivo fim, foram aplicadas a partir da participação e das contribuições da ciência da computação e do uso de recursos computacionais.

### Modelos matemáticos

Nesse ponto, introduz-se a ideia de modelos, no caso, modelos matemáticos e/ou computacionais, que são, de uma maneira ampla, representações de problemas do “mundo real” com o uso de técnicas matemáticas e/ou computacionais.

Acrescenta-se a essa ideia o fato de que esses modelos podem evoluir ao longo dos anos, com o avanço das técnicas matemáticas e computacionais e, se esses avanços, puderem ser sistematizados e aplicados às “necessidades e desejos” da área específica do estudo – por exemplo, genômica ou melhoramento genético

– então podem surgir novas formas de se buscar o avanço nessas áreas.

Em específico, a seleção genômica é exemplo de uma ferramenta para o melhoramento genético, que faz avançar meios e conhecimento para a seleção de animais.

A seleção genômica é realizada por meio de milhares de marcadores distribuídos ao longo do genoma dos animais avaliados, para análise e investigação do seu potencial genético, antes de que possam expressar seus fenótipos. Isto é, antes que seja possível mensurar seus atributos relacionados com a característica fenotípica que se quer avaliar.

Com o desenvolvimento do conhecimento científico e da tecnologia, em torno da genômica e da computação, tornou-se possível

a identificação de conjuntos com milhares de marcadores moleculares do tipo SNP<sup>1</sup>, sendo que, cada conjunto de marcadores moleculares, pode estar associado a diferentes características fenotípicas.

Sob o ponto de vista da genômica, algumas das dificuldades para a avaliação e seleção genômica podem ser a identificação dos elementos dos conjuntos de marcadores – isto é, quais são os marcadores relacionados com a característica fenotípica de interesse – e a dimensão dos conjuntos de marcadores – visto que, atualmente, pode-se chegar a quase 800 mil marcadores para cada animal a ser avaliado.

### Tratamento dos dados

As dificuldades no aspecto da genômica deixam claro que, no aspecto da computação, o tratamento dos dados, por si, já é um fator de dificuldade, pois, as informações sobre os marcadores moleculares dos indivíduos avaliados, podem tratar-se de arquivos de dados da ordem de terabytes.

Esses dois aspectos de tratamento do problema da avaliação e seleção genômica, remetem ao que foi abordado há pouco, onde conceitos e técnicas próprias genômica só

<sup>1</sup> SNP, do inglês, *single nucleotide polymorphism* ou polimorfismo de base única.

podem ser aplicadas – e conseqüentemente, chegar-se ao resultado esperado – a partir do desenvolvimento e uso adequado de métodos e estratégias da computação.

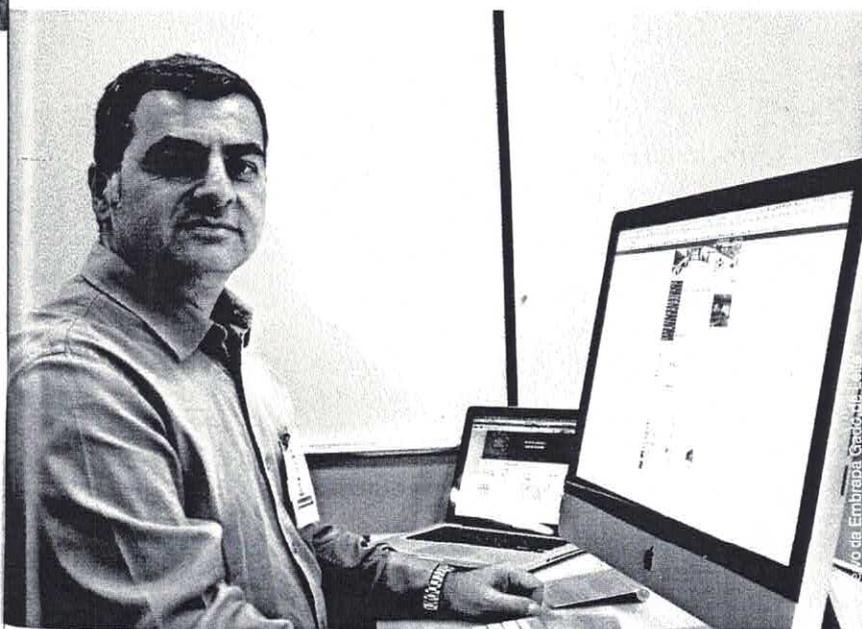
### Computação científica e genômica

A Embrapa realizou o TACG – Talking About Computing and Genomics<sup>2</sup>, um evento que reuniu as equipes técnicas de dois projetos que tratam de genômica animal e de modelos computacionais para bioinformática e biologia computacional, coordenados pela Empresa Brasileira de Pesquisa Agropecuária (Embrapa) e desenvolvidos em parceria com dezenas de instituições de ensino e pesquisa e instituições do setor produtivo voltadas para o agronegócio.

Como o próprio nome do evento sugere, o TACG permitiu uma ampla discussão de trabalhos em desenvolvimento em diferentes estágios onde pesquisadores das áreas envolvidas conheceram como a genômica pode avançar com o uso de novos recursos computacionais e, da mesma forma, para onde a computação deve voltar suas pesquisas para conseguir responder às expectativas.

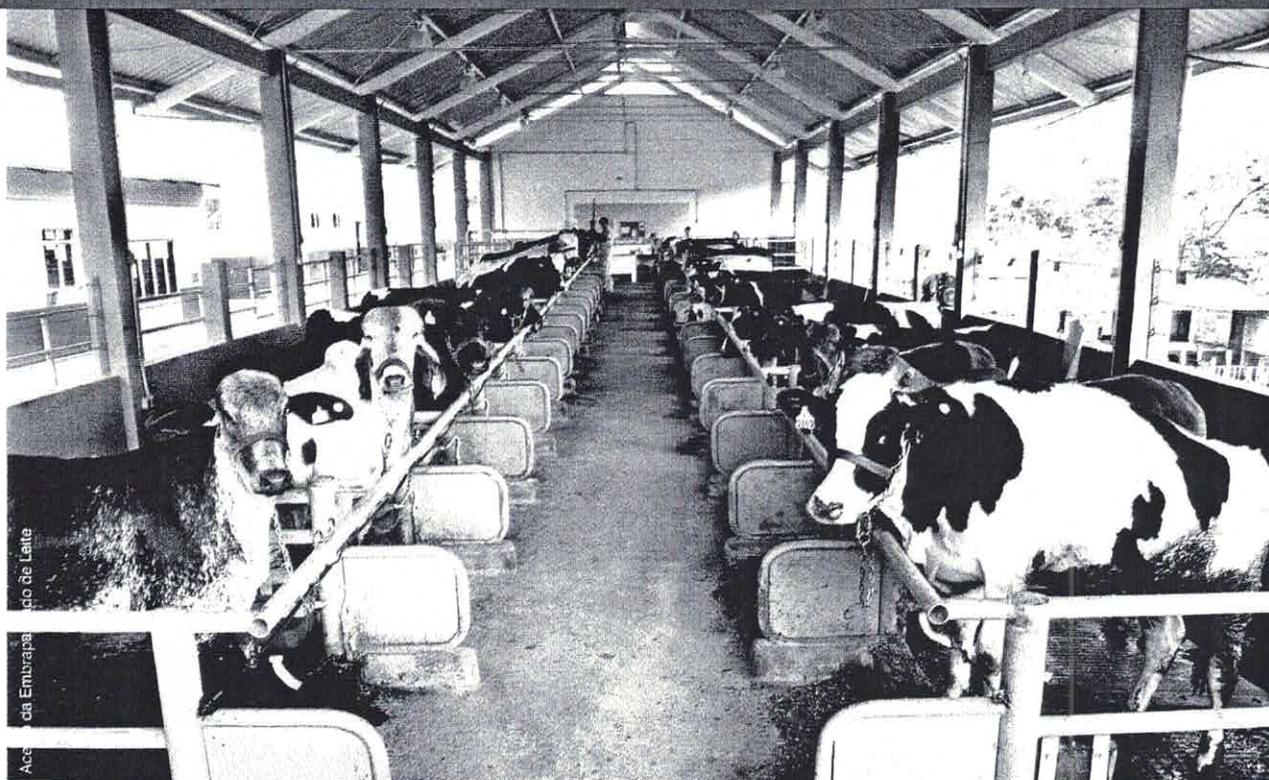
Os trabalhos do evento chegaram a ser surpreendentes e deixaram claro como a genômica “clássica” pode beneficiar-se de métodos compu-

<sup>2</sup> Maiores informações em <<http://www.cnpgl.embrapa.br/tacg/>>.



Wagner Arbex

“ Há vários anos são utilizados marcadores moleculares para a identificação de “locais” nas sequências genéticas, onde se encontram as informações sobre as características fenotípicas ”



Animais do rebanho da Embrapa Gado de Leite no Complexo Multiusuário de Bioeficiência e Sustentabilidade da Pecuária

tacionais e matemáticos para avançar na direção da fronteira do conhecimento em biotecnologia.

## SNPs

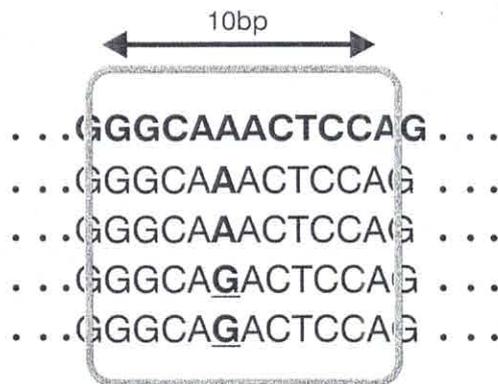
Há vários anos são utilizados marcadores moleculares para a identificação de "locais" nas sequências genéticas onde se encontram as informações sobre as características fenotípicas.

Entretanto, com a computação e a bioinformática, uma nova perspectiva foi aberta, frente ao volume de informações passíveis de serem tratados, assim como, às novas ferramentas e modelos computacionais que estão sendo utilizados.

Nesse caso, destacam-se a identificação dos chamados "snips", isto é, polimorfismos de base única, do inglês single nucleotide polymorphisms (SNPs), que são modificações de um único nucleotídeo, em uma dada sequência, quando comparada a outra, como exemplificado na ilustração ao lado.

Ou seja, SNPs são pares de bases, ou base pairs (bp), em uma única posição no DNA genômico, que se apresentam com diferentes alternativas nas sequências e podem ser encontrados no genoma de indivíduos normais em algumas populações ou grupos de indivíduos.

O que difere um indivíduo dos demais da sua espécie é o código genético, que, em sua essência, são as sequências de nucleotídeos que for-



Exemplo hipotético de um SNP, em uma "janela" com tamanho de 10 bp.

mam as moléculas e sequências de DNA, RNA e proteínas, que, por sua vez, interagem e formam as células, as quais também, por sua vez, interagem e formam os tecidos, os órgãos, até que, finalmente, formam os indivíduos.

Ou seja, no código genético, as diferenças se iniciam na ordem em que os nucleotídeos se apresentam para, posteriormente, após um complexo processo, originarem as proteínas. Essa é a importância dos SNPs, pois, em síntese, a alteração de um único nucleotídeo, uma única base, em uma dada sequência, pode alterar a produção de uma certa proteína e, se for o caso, o conjunto dessas alterações pode provocar variações nas características dos indivíduos da espécie.

Ano 06 - Número 12 - 2013 - R\$ 12,00 - www.animalbusinessbrasil.com.br

# animal *business* BRASIL

## MATO GROSSO

oportunidades  
de investimento  
na produção  
de carne bovina

CAMARÃO  
BRANCO  
produção  
high-tech

OVO DE  
GALINHA  
o mito do  
colesterol

MERCADO  
CHINÊS  
ascensão do  
poder de compra

0.0012  
ISSN 2227-132X  
9 772237 132007  
R\$ 12,00

