



## Delimitação de experimentos em genética genômica

Guilherme Jordão de Magalhães Rosa

Department of Dairy Science, University of Wisconsin, Madison – WI, USA 53705. E-mail: grosa@wisc.edu

**RESUMO** – Genética genômica é um termo utilizado para representar o estudo de processos genéticos controladores de caracteres fenotípicos de herança complexa, a partir da análise conjunta de informação relativa a fenótipos, estruturas de parentesco, marcadores moleculares e expressão gênica. Estudos de genética genômica são utilizados, por exemplo, para a estimação da herdabilidade de níveis de transcrição, para o mapeamento de locos controladores da expressão gênica (eQTL, do inglês *expression Quantitative Trait Loci*), e para o estudo de redes regulatórias. Genética genômica geralmente envolve experimentos com *microarrays*, os quais são ainda bastante caros e trabalhosos, limitando o tamanho amostral e conseqüentemente o poder estatístico de tais estudos. Desta maneira, é essencial que tais experimentos sejam otimizados do ponto de vista do delineamento, a partir de criteriosa escolha das amostras (indivíduos) a serem utilizadas, e do controle rigoroso dos vários fatores que podem afetar as variáveis-resposta de interesse. Outro ponto fundamental na condução de tais experimentos refere-se à marcação das amostras de mRNA com os fluoróforos e ao pareamento das mesmas em cada lâmina de *microarray*, os quais devem ser cuidadosamente planejados para que não haja confundimento entre estes efeitos e os fatores biológicos de interesse. Nesta apresentação serão discutidas algumas estratégias para o planejamento de estudos de genética genômica, incluindo a seleção de indivíduos objetivando-se a maximização da dissimilaridade genética ou do número de eventos de recombinação, bem como a condução eficiente dos ensaios com *microarrays* para diferentes objetivos experimentais.

**Palavras-chave:** delineamentos ótimos, eQTL, fenotipagem seletiva, genética genômica, *microarray*

## Experimental design in genetical genomics

**ABSTRACT** – Genetical genomics experiments combine information on phenotypic traits, molecular markers and gene expression to study the genetic mechanisms governing variation in complex traits. Such studies can be used, for example, to estimate heritabilities of mRNA transcript abundances, to map expression quantitative trait loci (eQTL), and to infer regulatory gene networks. Microarray experiments, however, can be extremely costly and time consuming, which may limit sample sizes and statistical power. Thus it is crucial to optimize experimental designs by carefully choosing the subjects to be assayed, and by cautiously controlling systematic factors affecting the system. Also, a rigorous strategy should be used for allocating mRNA samples across slides and dye labeling, so that effects of interest are not confounded with nuisance factors. In this presentation, we review some designs strategies for genetical genomics studies, including the selection of individuals for increased genetic dissimilarity and for a higher number of recombination events, as well as efficient microarray experiment layouts for various experimental goals.

**Key Words:** eQTL, gene expression, optimal design, selective phenotyping, transcriptional profiling;

### Introdução

Genética genômica e genômica quantitativa são termos utilizados para representar o estudo de processos genéticos controladores de caracteres fenotípicos de herança complexa, a partir da análise conjunta de informação relativa a fenótipos, estruturas de parentesco, marcadores

moleculares e expressão gênica (Jansen & Nap, 2001; Darvasi, 2003; Pomp *et al.*, 2004). Por exemplo, experimentos com *microarrays* (ou microarranjos de DNA) têm sido utilizados para a avaliação dos padrões de expressão gênica em indivíduos com informação genotípica de marcadores moleculares. Valores de expressão dos diferentes genes são então tratados como caracte-

rísticas fenotípicas múltiplas de maneira que procedimentos de mapeamento de locos controladores de caracteres fenotípicos quantitativos (ou QTL, do termo em inglês *Quantitative Trait Loci*) são utilizados para a detecção de polimorfismos associados com a variabilidade nos padrões de expressão, os quais são geralmente denominados por eQTL. O efeito de variações alélicas na expressão gênica na mesma região cromossômica é denominado efeito *cis*, enquanto que o efeito de polimorfismos na expressão de genes em outras regiões do genoma é denominado efeito *trans* (Jansen & Nap, 2001).

Este tipo de metodologia tem sido utilizado para a identificação de genes candidatos (Schadt *et al.*, 2003; Bystrykh *et al.*, 2005; Hubner *et al.*, 2005), para o estudo de redes regulatórias (Brem *et al.*, 2002; Schadt *et al.*, 2003; Yvert *et al.*, 2003) e, principalmente, para a identificação de regiões cromossômicas controladoras da expressão de múltiplos genes, denominadas como regiões *hot spots* (Brem *et al.*, 2002; Schadt *et al.*, 2003; Morley *et al.*, 2004; Bystrykh *et al.*, 2005; Chesler *et al.*, 2005; Hubner *et al.*, 2005). Estas regiões cromossômicas referem-se a fatores reguladores associados à transcrição de genes relacionados à determinadas funções as quais podem ser de interesse, por exemplo, para tratamento terapêutico genético (Cox, 2004; Schadt *et al.*, 2003). Além do mapeamento de eQTL e da detecção de regiões *hot spots*, experimentos de genética genômica têm sido utilizados também no estudo do efeito *trans* de variações alélicas de determinados locos, tais como genes candidatos ou transgenes, na expressão de outros genes do genoma (Noueiry *et al.*, 2000), e na estimação da herdabilidade dos níveis de transcrição utilizando-se informação de indivíduos aparentados (Gibson *et al.*, 2004; Monks *et al.*, 2004; Dermitzakis & Stranger, 2006). Para uma revisão de diferentes métodos estatísticos utilizados em genética genômica consulte, por exemplo, Alberts *et al.* (2005), Carlborg *et al.* (2005), Kendzioriski & Wang (2006), e Rosa *et al.* (2006b).

Como estudos de genética genômica envolvem técnicas laboratoriais caras e trabalhosas, como por exemplo a genotipagem de grande número de marcadores moleculares como polimorfismos de base única (SNP, do inglês *Single Nucleotide Polimorphisms*) e experimentos de expressão

gênica utilizando-se a tecnologia de *microarrays* ou a técnica da transcrição reversa quantitativa acoplada com a reação de polimerase em cadeia (qRT-PCR, do inglês *quantitative Reverse Transcription Polymerase Chain Reaction*), é imprescindível um cuidadoso planejamento experimental de tais ensaios. Nesta apresentação serão discutidos alguns princípios básicos de delineamento de experimentos com *microarrays* e as estratégias experimentais comumente utilizadas em estudos de genética genômica para diferentes objetivos experimentais, tais como a comparação de padrões de expressão gênica de diferentes grupos genotípicos para genes candidatos, o mapeamento de eQTL, e a estimação de herdabilidades dos níveis de transcrição gênica.

### Planejamento de experimentos em genética genômica

O planejamento de estudos de genética genômica envolve uma série de aspectos. Da mesma maneira que qualquer outro estudo de mapeamento de QTL, estudos de genética genômica requerem, por exemplo, a escolha das raças ou linhagens, bem como a estratégia de cruzamentos a ser utilizada, como retrocruzamento, F<sub>2</sub>, etc. Além disso, estudos de expressão gênica envolvem também a escolha cuidadosa do(s) tecido(s) ou tipo(s) de célula(s), bem como o(s) estágio(s) de desenvolvimento a serem amostrados. Especificamente em relação aos experimentos com *microarrays*, tem-se ainda que decidir que tipo específico de tecnologia de *microarray* mais se adequa aos objetivos experimentais em questão, bem como quantas lâminas são necessárias para que desejada eficiência experimental seja alcançada. Além disso, tem-se que selecionar os indivíduos a serem amostrados e, no caso de experimentos com *microarrays* no sistema de hibridizações competitivas, tem-se que planejar como as amostras serão pareadas e marcadas em cada hibridização (Rosa *et al.*, 2007).

A escolha da tecnologia de *microarray* a ser utilizada é muitas vezes limitada pela disponibilidade de alternativas para a espécie animal a ser estudada. Por exemplo, *microarrays* com sondas de cDNA podem ser construídos somente para espécies das quais etiquetas de seqüências

expressas (EST, na sigla em inglês para *Expressed Sequence Tags*) já foram obtidas de bibliotecas de cDNA. Já *microarrays* que utilizam sondas de oligonucleotídeos somente são possíveis se seqüências de DNA estão disponíveis e, ainda, a obtenção de um grupo eficiente de sondas só é possível para espécies com o genoma completamente seqüenciado. Em relação ao número de lâminas a serem utilizadas num experimento específico, o mesmo é em geral determinado por restrições orçamentárias, exceto talvez em estudos relativos à pesquisa médica. Desta maneira, nesta apresentação serão discutidos dois aspectos relativos ao planejamento de experimentos em genética genômica, os quais referem-se a conceitos de estatística e delineamento experimental:

- 1) Fenotipagem seletiva, a qual se refere à sub-amostragem dos indivíduos a serem utilizados nos experimentos com *microarrays*, e
- 2) Esquema experimental dos ensaios com *microarrays*.

### Fenotipagem seletiva

Até alguns anos atrás o custo de genotipagem era um fator limitante em estudos de mapeamento. Eram utilizadas então estratégias experimentais tais como a genotipagem seletiva (Lebowitz *et al.*, 1987; Lander & Botstein, 1989; Darvasi & Soller, 1992; Allison *et al.*, 1998). Em estudos de genética genômica, entretanto, a fenotipagem é o fator limitante dados os custos ainda bastante altos dos experimentos com *microarrays*. Desta maneira, a avaliação dos padrões de expressão gênica de uma sub-amostra dos animais disponíveis no estudo passa a ser uma estratégia natural para a redução dos custos dos estudos de genética genômica.

Uma estratégia de fenotipagem seletiva foi proposta por Jin *et al.* (2004), a qual utiliza informação de marcadores moleculares para a sub-amostragem de indivíduos com o objetivo de se maximizar a dissimilaridade genética na amostra selecionada, e conseqüentemente aumentar o poder de detecção de QTL. O procedimento proposto por estes autores compara o genótipo dos marcadores de todos os indivíduos disponíveis e genotipados, e utiliza um algoritmo de otimização para a seleção de uma sub-amostra de indivíduos a serem fenotipados, *i.e.*, a serem amostrados para o experimento de expressão gênica com

*microarrays*. A medida de similaridade genética utilizada por Jin *et al.* (2004) refere-se ao número de alelos em comum entre dois indivíduos (0, 1 ou 2), para cada um dos marcadores genéticos, sendo que a escolha dos indivíduos pode ser baseada na informação de todos os marcadores disponíveis ou, alternativamente, se restringir a marcadores referentes a regiões cromossômicas previamente detectadas como importantes para as características fenotípicas de interesse.

Jin *et al.* (2004) utilizaram um estudo de simulação para comparar a eficiência de detecção de QTL desta estratégia de fenotipagem seletiva relativamente à uma amostragem aleatória do grupo de indivíduos disponíveis. As simulações consideraram experimentos com indivíduos  $F_2$ , um único cromossomo com marcadores a cada 10 cM, e um único QTL com valores de herdabilidade  $h^2 = 0,25; 0,50$  e  $0,75$ . Foram ainda estudados diferentes cenários em termos do número de indivíduos disponíveis ( $N = 50 - 200$ ), do número de indivíduos selecionados ( $n = 10 - N$ ), e proporções de indivíduos selecionados (10–90%). Os resultados demonstraram que, para um determinado tamanho de sub-amostra ( $n = 50$ ), a sensibilidade (*i.e.*, a porcentagem de simulações nas quais o QTL foi detectado) aumentou com o aumento do número total de indivíduos disponíveis até quando a proporção de indivíduos selecionados atingiu aproximadamente 50% (para  $h^2 = 0,25$ ). Numa situação com um número fixo de indivíduos disponíveis ( $N = 100$ ), incrementos na proporção de indivíduos selecionados propiciaram aumentos maiores de sensibilidade na fenotipagem seletiva baseada na estratégia de dissimilaridade genética do que na amostragem aleatória. Mais uma vez, entretanto, não houve aumento relevante em sensibilidade quando mais do que 50% dos indivíduos foram selecionados, especialmente para situações com alta herdabilidade, indicando que a maior parte da informação necessária para a detecção de QTL foi recuperada com 50% de fenotipagem seletiva.

A medida de similaridade genética considerada por Jin *et al.* (2004) tende a selecionar indivíduos que são predominantemente homocigotos para alelos distintos nos marcadores moleculares considerados. Por exemplo, numa população  $F_2$  obtida a partir do cruzamento entre linhagens consanguíneas ou entre raças puras, tem-se propor-

ções esperadas de 1:2:1 para os genótipos A, H e B, onde A e B representam indivíduos homocigotos para alelos das linhagens A e B, respectivamente, e H representa indivíduos heterocigotos. A metodologia de Jin *et al.* (2004), quando utilizada neste tipo de situação, tende a selecionar indivíduos de maneira que a proporção 1:0:1 é favorecida. Este procedimento, entretanto, é recomendado somente se o interesse central do experimento for relacionado à detecção de efeitos genéticos aditivos. Efeitos aditivos, no entanto, ainda que muito importantes e geralmente de interesse em tais estudos, são também os efeitos mais fáceis de serem detectados. Para a estimação de efeitos genéticos mais complexos, entretanto, outras combinações genotípicas são necessárias na sub-amostra estudada (Rosa *et al.*, 2006a).

Por exemplo, Piepho (2005) propôs um critério alternativo de fenotipagem seletiva, favorecendo a proporção 1:2:1 para os genótipos A:H:B na sub-amostra, a qual maximiza a eficiência na estimação de efeitos de dominância. Uma generalização deste procedimento foi proposta por Bueno *et al.* (2006), na qual a proporção alvo para os genótipos A:H:B depende dos objetivos do experimento. Se o objetivo central de um experimento com uma população  $F_2$  refere-se somente à estimação de efeitos aditivos ou de efeitos de dominância, a sub-amostragem proposta por Bueno *et al.* (2006) coincidirá com as de Jin *et al.* (2004) e de Piepho (2005), respectivamente. No entanto, se tanto efeitos aditivos quanto efeitos de dominância forem de interesse, proporções intermediárias podem ser mais interessantes. Por exemplo, se a mesma ênfase for colocada na estimação de efeitos aditivos e de dominância a proporção ótima dos genótipos A:H:B será 0,293:0,414:0,293.

Os métodos propostos por Jin *et al.* (2004), Piepho (2005) e Bueno *et al.* (2006) maximizam o poder de detecção de QTL e eficiência na estimação de efeitos genéticos específicos. Estes procedimentos, no entanto, não objetivam ou não maximizam a eficiência na estimação da posição dos QTLs. Neste contexto, de Leon & Rosa (2005) propuseram um critério de fenotipagem seletiva para a maximização do número de recombinações observadas na sub-amostra selecionada. Num estudo de simulação com população referente à retrocruzamento, um único cromossomo com 10 marcadores equidistantes e um QTL, e diferentes

números de indivíduos disponíveis e sub-amostrados, os autores concluíram que a fenotipagem seletiva baseada no número de recombinações melhorou sensivelmente a precisão das estimativas das posições dos QTLs, mesmo quando comparada com seleção baseada no conceito de dissimilaridade genética aplicada a marcadores na vizinhança do QTL. Os autores mostraram, também, que não houve perda significativa no poder de detecção ou na eficiência de estimação dos efeitos dos QTLs.

Resultados similares foram apresentados por Jannink (2005) e Xu *et al.* (2005), os quais desenvolveram simulações em contextos ainda mais gerais, com diferentes comprimentos cromossômicos, espaçamento variável entre os marcadores, e diferentes números de QTLs. As metodologias propostas por Jannink (2005) and Xu *et al.* (2005), inclusive, são mais gerais do que a apresentada por de Leon & Rosa (2005) pois as mesmas favorecem não somente um maior número de recombinações, mas também uma distribuição mais uniforme das recombinações ao longo do genoma.

Um outro objetivo experimental também considerado em estudos de genética genômica refere-se à estimação de componentes de variância e herdabilidade associados a expressão gênica (Gibson *et al.*, 2004; Monks *et al.*, 2004). Nestas circunstâncias, um critério eficiente para fenotipagem seletiva deve considerar o parentesco entre os indivíduos disponíveis, como discutido por Bueno *et al.* (2006). Estes autores implementaram um algoritmo para situações desta natureza e apresentaram alguns exemplos envolvendo grupos de meio-irmãos, irmãos completos, e estruturas de parentesco mais complexas, mostrando que a estratégia de seleção é bastante específica para cada estrutura de parentesco considerada.

### **Esquema experimental para ensaios com *microarrays***

Após a seleção dos indivíduos a serem utilizados nos experimentos com *microarrays*, outro importante passo no delineamento experimental se faz necessário, especialmente se for utilizada a tecnologia de hibridizações competitivas. Nestes casos, tem-se ainda que se decidir como as diferentes amostras serão pareadas e marcadas para cada

hibridização (Wit *et al.*, 2005; Bueno *et al.*, 2006; Rosa *et al.*, 2007).

Os esquemas experimentais mais comumente utilizados em experimentos com *microarrays* de hibridizações competitivas referem-se às estruturas referentes aos delineamentos com amostra referência e delineamentos circulares (Kerr & Churchill, 2001; Yang & Speed, 2002). Uma revisão destes dois tipos básicos de delineamentos para *microarrays* bem como uma discussão das vantagens e desvantagens de cada um pode ser encontrado, por exemplo, em Rosa *et al.* (2007) neste mesmo exemplar da **Revista Brasileira de Zootecnia**. A maioria da literatura existente na área de delineamento de experimentos com *microarrays*, entretanto, refere-se a situações nas quais não existe um componente genético definindo os diferentes grupos experimentais, de maneira que os resultados não se aplicam diretamente a um contexto de genética genômica. Somente mais recentemente é que alguns estudos propondo estratégias experimentais mais eficientes para genética genômica foram apresentados, os quais são discutidos a seguir.

Fu & Jansen (2006), por exemplo, propuseram uma estratégia de delineamento para o pareamento de amostras referentes a linhagens recombinantes (RIL, de *Recombinant Inbred Lines*) para experimentos com hibridizações competitivas. O procedimento por eles sugerido, denominado *delineamento com pareamentos distantes*, baseia-se em dois princípios básicos. Primeiro, para um determinado número de lâminas, é geralmente mais eficiente o aumento do número de replicações biológicas do que replicações técnicas; segundo, amostras devem ser pareadas de maneira que a dissimilaridade dentro das lâminas de *microarray* deve ser maximizada, relativamente à dissimilaridade entre lâminas. É interessante se notar que o procedimento de pareamento de amostras proposto por Fu & Jansen (2006) pode ser utilizado em conjunção com alguma estratégia de fenotipagem seletiva, como por exemplo os métodos de Jin *et al.* (2004), Jannink (2005) ou Xu *et al.* (2005), em situações nas quais o número de lâminas for menor do que o número de indivíduos disponíveis.

O delineamento com pareamentos distantes, no entanto, é recomendado somente em situações nas quais o interesse central do experimento refere-se ao estudo de efeitos aditivos, ou em estudos com

populações nas quais somente dois genótipos são possíveis para cada loco, como por exemplo, em retrocruzamentos, RIL ou populações haplodiplóides. Se outros efeitos forem também de interesse, outras alternativas de pareamento podem ser necessárias. Neste contexto, Piepho (2005) discutiu delineamentos eficientes para experimentos com *microarrays* em sistema de hibridizações competitivas quando o interesse principal refere-se à estimação de efeitos de dominância. Nestes casos, como demonstrado, pareamentos envolvendo uma amostra homozigota e outra heterozigota é mais eficiente do que pareamentos com duas amostras homozigotas (ainda que para alelos distintos) ou duas amostras heterozigotas.

O procedimento proposto por Piepho (2005), no entanto, aplica-se somente a situações com duas linhagens recombinantes e o cruzamento F1, ou a situações com um único gene candidato bialélico (Rosa *et al.*, 2006a). Uma metodologia mais geral para o pareamento de amostras em experimentos com *microarrays* em estudos de genética genômica (incluindo também a otimização da marcação das amostras em cada par) foi discutido por Bueno *et al.* (2006).

Estes autores demonstraram que as amostras devem ser pareadas e marcadas de maneira a favorecer hibridizações referentes a comparações mais informativas em termos dos parâmetros genéticos de interesse. Os resultados apresentados utilizando-se exemplos com múltiplos locos generalizam o conceito de pareamentos distantes, bem como a estratégia de pareamento de Piepho para a estimação de efeitos de dominância. Como demonstrado, os delineamentos ótimos dependem dos efeitos genéticos de interesse, e de como estes efeitos são ponderados no critério de otimização.

Por exemplo, se o interesse se restringe aos efeitos aditivos, o delineamento ótimo refere-se ao de pareamentos distantes. Se o interesse, no entanto, for principalmente em relação aos efeitos de dominância, o delineamento resultante vai tender a parear homozigotos com heterozigotos, como proposto por Piepho (2005). Entretanto, se interações epistáticas são também consideradas, a estrutura de delineamento torna-se mais complexa. Em tais situações, o pareamento das amostras deve considerar a combinação genotípica para os múltiplos locos.

Uma outra generalização do conceito de

pareamentos distantes proposto por Fu & Jansen (2006) foi também apresentada por Bueno *et al* (2006) para experimentos objetivando-se a estimação de componentes de variância e herdabilidade dos padrões de expressão gênica. Como demonstrado, em experimentos desta natureza considerando indivíduos com determinada estrutura de parentesco, delineamentos eficientes tendem a parear amostras referentes a indivíduos menos aparentados.

### Considerações finais

Nesta apresentação são discutidas estratégias de delineamento para experimentos com *microarray* em estudos de genética genômica com diferentes objetivos experimentais, tais como a comparação dos padrões de expressão gênica de diferentes grupos genotípicos, o mapeamento de eQTL, ou a estimação de herdabilidades dos padrões de expressão. O planejamento de um delineamento experimental eficiente geralmente envolve uma sub-amostragem dos indivíduos disponíveis (denominada fenotipagem seletiva), para o posterior pareamento e marcação das amostras (referente às hibridizações competitivas em cada lâmina de *microarray*).

As diferentes estratégias de fenotipagem seletiva dependem somente do material biológico disponível e dos objetivos do experimento, de maneira que elas são as mesmas para qualquer tipo de tecnologia de *microarray* a ser utilizada, como a de hibridizações independentes (*e.g.* slide Affymetrix) ou de hibridizações competitivas (sistema de duas cores utilizando sondas de cDNA ou de seqüências oligonucleotídeas). A maneira como as amostras são organizadas para as hibridizações, no entanto, dependem também do tipo de tecnologia a ser utilizada. No caso de *microarrays* com hibridizações independentes, como cada lâmina é utilizada para uma única amostra, o experimento é bastante simples. Já no caso de experimentos com hibridizações competitivas, existem sempre inúmeras alternativas para o pareamento e marcação das amostras. A mais simples alternativa refere-se ao delineamento com amostra referência, o qual se assemelha a um experimento com hibridizações independentes, uma vez que cada amostra de interesse é hibridizada numa lâmina independente (a única

diferença é que neste caso cada lâmina recebe também a amostra referência). Entretanto, geralmente têm-se outras estruturas de delineamento que propiciam maior eficiência e precisão para tais experimentos, como discutido por Rosa *et al.* (2007).

Em geral, replicação biológica é mais vantajosa do que replicação técnica do ponto de vista estatístico. Por exemplo, se um delineamento com amostra referência com  $2 \times n$  lâminas de *microarray* é considerado, uma maior precisão é alcançada se  $2 \times n$  indivíduos são considerados, um em cada lâmina, do que se somente  $n$  indivíduos são utilizados, cada um com amostras hibridizados em duas lâminas (*i.e.* replicação técnica) com marcação invertida. Como discutido nesta apresentação, experimentos ainda mais eficientes podem ser obtidos num contexto de delineamentos ótimos, buscando delineamentos eficientes para determinados objetivos experimentais. Diferentes combinações de replicação técnica e biológica podem até ser utilizadas, mas a análise estatística dos dados obtidos deve considerar a estrutura de delineamento utilizada (Rosa *et al.*, 2005). Outra estratégia comumente utilizada em experimentos com *microarrays* refere-se à mistura de amostras de mRNA na tentativa de se reduzir a variabilidade biológica (Churchill, 2002). Em estudos de genética genômica, entretanto, a mistura de amostras não é aconselhada, exceto em situações envolvendo indivíduos geneticamente idênticos, como em populações de RIL.

Nesta apresentação foram discutidas estratégias de fenotipagem seletiva baseadas somente em informação de marcadores moleculares ou de parentesco entre indivíduos. Alternativamente, a fenotipagem seletiva pode também envolver informação referente a valores fenotípicos para características correlacionadas (Medugorac & Soller, 2001), ou uma combinação de informação referente a fenótipo e marcadores moleculares (Wang & Nettleton, 2006), ou fenótipo e estruturas de família (Rosa *et al.*, 2006b; Cardoso *et al.*, 2007). As vantagens e desvantagens de cada método dependem dos objetivos do experimento em questão, e nas suposições dos modelos estatísticos a serem utilizados na análise dos dados. Vale ressaltar, entretanto, que geralmente o mecanismo de sub-amostragem utilizado deve ser considerado na análise dos dados observados.

### Agradecimento

Este projeto foi parcialmente financiado pelo auxílio 2004-33120-15204 do *United States Department of Agriculture* (USDA).

### Literatura citada

- ALBERTS, R.; FU, J.; SWERTZ, M.A. *et al.* Combining microarrays and genetic analysis. **Briefings in Bioinformatics**, v.6, n.2, p.135-145, 2005.
- ALLISON, D.B.; HEO, M.; SCHORK, N.J. *et al.* Extreme selection strategies in gene mapping studies of oligogenic quantitative traits do not always increase power. **Human Heredity**, v.48, n.2, p.97-107, 1998.
- BREM, R.B.; YVERT, G.; CLINTON, R. *et al.* Genetic dissection of transcriptional regulation in budding yeast. **Science**, v.296, p.752-755, 2002.
- BUENO, J.S.D.; GILMOUR, S.G.; ROSA, G.J.M. Design of microarray experiments for genetical genomics studies. **Genetics**, v.74, n.2, p.945-957, 2006.
- BYSTRYKH, L.; WEERSING, E.; DONTJE, B. *et al.* Uncovering regulatory pathways affecting hematopoietic stem cell function using "genetical genomics." **Nature Genetics**, v.37, p.225-232, 2005.
- CARDOSO, F.F.; STEIBEL, J.P.; ROSA, G.J.M. *et al.* Assessment of different selective phenotyping and data analysis strategies for genetical genomics studies with outbred F2 populations. In: ADSA/PSA/AMPA/ASAS 2007 JOINT ANNUAL MEETING, San Antonio, Texas, USA, 8-12 de julho, 2007.
- CARLBORG, O.; DE KONING, D.J.; MANLY, K.F. *et al.* Methodological aspects of the genetic dissection of gene expression. **Bioinformatics**, v.21, n.10, p.2383-2393, 2005.
- CHESLER, E.J.; LU, L.; SHOU, S. *et al.* Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. **Nature Genetics**, v.37, p.233-242, 2005.
- CHURCHILL, G.A. Fundamentals of experimental design for cDNA microarrays. **Nature Genetics**, v.32, p.490-495, 2002.
- COX, N.J. Human genetics - An expression of interest. **Nature Genetics**, v.430, p.733-734, 2004
- DARVASI, A. Gene expression meets genetics. **Nature**, v.422, p.269-71, 2003.
- DARVASI, A.; SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. **Theoretical and Applied Genetics**, v.85, n.2-3, p.353-359, 1992.
- DE LEON, N.; ROSA, G.J.M. Optimization of selective phenotyping for QTL mapping. In: PLANT AND ANIMAL GENOME XIII MEETINGS, P860, p.283, 2005. ([http://www.intl-pag.org/13/abstracts/PAG13\\_P860.html](http://www.intl-pag.org/13/abstracts/PAG13_P860.html))
- DERMITZAKIS, E.T.; STRANGER, B.E. Genetic variation in human gene expression. **Mammalian Genome**, v.17, n.6, p.503-508, 2006.
- FU, J.; JANSEN, R.C. Optimal design and analysis of genetic studies on gene expression. **Genetics**, v.172, p.1993-1999, 2006.
- GIBSON, G.; RILEY-BERGER, R.; HARSHMAN, L. *et al.* Extensive sex-specific non-additivity of gene expression in *Drosophila melanogaster*. **Genetics**, v.167, p.1791-1799, 2004.
- HUBNER, N.; WALLACE, C.A.; ZIMBAHL, H. *et al.* Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. **Nature Genetics**, v.37, p.243-253, 2005.
- JANNINK JL. Selective phenotyping to accurately map quantitative trait loci. **Crop Science**, v.45, n.3, p.901-908, 2005.
- JANSEN, R.C.; NAP, J. Genetical genomics: the added value from segregation. **Trends in Genetics**, v.17, p.388-91, 2001.
- JIN, C.F.; LAN, H.; ATTIE, A.D. *et al.* Selective phenotyping for increased efficiency in genetic mapping studies. **Genetics**, v.168, n.4, p.2285-2293, 2004.
- KENDZIORSKI, C.; WANG, P. A review of statistical methods for expression quantitative trait loci mapping. **Mammalian Genome**, v.17, p.509-517, 2006.
- KERR, M.K.; CHURCHILL, G.A. Experimental design for gene expression microarrays. **Biostatistics**, v.2, p.183-201, 2001.
- LANDER, E.S.; BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v.121, n.1, p.185-199, 1989.
- LEBOWITZ, R.J.; SOLLER, M.; BECHMAN, J.S. Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. **Theoretical and Applied Genetics**, v.73, p.556-562, 1987.
- MEDUGORAC, I.; SOLLER, M. Selective genotyping with a main and a correlated trait. **Journal of Animal Breeding and Genetics**, v.118, p.285-295, 2001.
- MONKS, S.A.; LEONARDSON, A.; ZHU, H. *et al.* Genetic inheritance of gene expression in human cell lines. **American Journal of Human Genetics**, v.75, p.1094-1105, 2004.
- MORLEY, M.; MOLONY, C.M.; WEBER, T.M. *et al.* Genetic analysis of genome-wide variation in human gene expression. **Nature**, v.430, p.743-747, 2004.
- NOUEIRY, A.O.; CHEN, J.; AHLQUIST, P. A mutant allele of essential, general translation initiation factor DED1 selectively inhibits translation of a viral mRNA. **Proceeding of the National Academy of Science USA**, v.97, p.12985-12990, 2000.
- PIEPHO, H.P. Optimal allocation in designs for assessing heterosis from cDNA gene expression data. **Genetics**, v.171, p.359-364, 2005.
- POMP, D.; ALLAN, M.F.; WESOLOWSKI, S. Quantitative Genomics: Exploring the genetic architecture of complex trait predisposition. **Journal of Animal Science**, v.82, p.E300-E312, 2004.
- ROSA, G.J.M.; DE LEON, N.; ROSA, A.J.M. Review of microarray experimental design strategies for genetical genomics studies. **Physiological Genomics**, v.28, p.15-23, 2006a.
- ROSA, G.J.M.; ROCHA, L.B.; FURLAN, L.R. Estudos de expressão gênica utilizando-se *microarrays*: delineamento, análise, e aplicações na pesquisa zootécnica. **Revista Brasileira de Zootecnia**, 2007 (in press).
- ROSA, G.J.M.; STEIBEL, J.P.; TEMPELMAN, R.J. Reassessing design and analysis of two-colour microarray experiments using mixed effects models. **Comparative and Functional Genomics**, v.6, p.123-131, 2005.
- ROSA, G.J.M.; TEMPELMAN, R.J.; ERNST, C.W. *et al.* Combining molecular marker information and gene expression profiling for studying complex traits. In: 8<sup>TH</sup> WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION. Belo Horizonte, Brasil, 2006b. (<http://www.wcgalp8.org.br/>).
- SCHADT, E.E.; MONKS, S.A.; DRAKE, T.A. *et al.* Genetics of gene expression surveyed in maize, mouse and man. **Nature**, v.422, p.297-302, 2003.
- WANG, D.; NETTLETON, D. Identifying genes associated with a quantitative trait or quantitative trait locus via selective transcriptional profiling. **Biometrics**, v.62, n.2, p.504-514, 2006.
- WIT, E.; NOBILE, A.; KHANIN, R. Near-optimal designs for dual channel microarray studies. **Applied Statistics**, v.54, p.817-830, 2005.

- XU, Z.; ZOU, F.; VISION, T.J. Improving QTL mapping resolution in experimental crosses by the use of genotypically selected samples. **Genetics**, v.170, p.401-408, 2005.
- YANG, Y.H.; SPEED, T. Design issues for cDNA microarray experiments. **Nature Reviews Genetics**, v.3, p.579-588, 2002.
- YVERT, G.; BREM, R.B.; WHITTLE, J. *et al.* Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. **Nature Genetics**, v.35, p.57-64, 2003.