HAL
open science

A factor model to analyse heterogeneity in gene expression in a context of QTL characterization

Yuna Blum, Sandrine Lagarrigue, David Causeur

To cite this version:

Yuna Blum, Sandrine Lagarrigue, David Causeur. A factor model to analyse heterogeneity in gene expression in a context of QTL characterization. XXXII International Conference on Animal Genetics (ISAG), Jul 2010, Edinburgh, United Kingdom. 2010, Proceedings of the 32nd International Conference on Animal Genetics. hal-02754275

HAL Id: hal-02754275

https://hal.inrae.fr/hal-02754275

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

P2001 Alteration of gene expression in mammary gland tissue of dairy cows in response to dietary unsaturated fatty acids

Núria Mach[1], Antoon Jacobs[2], Leo Kruijt[1], Jurgen van Baal[2] and Mari Smits[1].
[1]Animal Breeding and Genomics Centre, Wageningen UR Livestock Research,
P.O. Box 135, 6700 AC Wageningen, The Netherlands.
[2]Animal Nutrition Group, Wageningen University, P.O. Box 338,
6700 AH Wageningen, The Netherlands.
Corresponding author: nuria.mach@wur.nl

The aim of this study was to determine the effects of unprotected dietary unsaturated fatty acids (UFA) from different plant oils on gene expression in the mammary gland of grazing dairy cows. A total of twenty-eight Holstein-Friesian dairy cows were randomly assigned to four concentrated UFA-sources for 23 days after which all 28 cows were switched to a non-UFA-supplemented concentrate for an additional 28 days. On the last day of both periods, mammary gland biopsies were taken to study genome-wide differences in gene expression on Affymettrix GeneChip® Bovine Genome Arrays. Supplementation with UFA increased milk yield and decreased milk fat and protein content. Furthermore, the proportion of *de novo* fatty acids in the milk was reduced whereas that of long-chain fatty acids increased. A total of 3,490 genes were found to be significantly affected by UFA supplementation. Gene sets related to cell growth, proliferation and development, apoptosis, cell cycle, signalling, nutrient metabolic process, as well as immune system response were predominantly down-regulated by UFA supplementation. In contrast, gene sets associated with electron transport chain, and oxidative phosphorylation processes were up-regulated. Therefore, supplementing grazing dairy cows with unprotected dietary UFA can improve the health and nutrition quality aspects of dairy milk, but may also affect mammary gland integrity and health.

P2002 Structural variation in the pig genome using next-generation sequencing

João Fadista, Bujie Zhan, Jakob Hedegaard, Christian Bendixen.
Faculty of Agricultural Sciences, Aarhus University, Blichers Allé 20, Postboks 50,
DK-8830, Tjele, Denmark.
Corresponding author: joao.fadista@agrsci.dk

Several recent studies have described copy number variation (CNV) and insertions/deletions (indels) as common and frequent structural variation present in mammalian genomes. Although the detection of such variation has been traditionally accomplished with SNP or CGH arrays, next-generation sequencing has recently proven to achieve better resolution and accuracy.

In this study, we report the detection of structural variation in the pig genome by analyzing whole-genome Illumina paired-end sequencing data of 4 Duroc boars at 15 fold coverage each. We describe genome-wide maps of CNV and segmental duplication based on read depth of coverage, while indels are detected in intra-reads and in inter-reads using paired-end information. Functional analysis provides genes and pathways putatively associated with important known QTLs.

This study provides the highest-resolution map of genomic structural variation throughout the porcine genome to date and reveals interesting targets for follow-up studies.

These results are obtained through the EC-funded FP6 Project "**SABRE**".

P2003 A factor model to analyze heterogeneity in gene expression in a context of QTL characterization

Yuna Blum[1,2,3], Sandrine Lagarrigue[1,2] and David Causeur[3].
[1]Agrocampus Ouest, UMR598, Animal Genetics, 35000 Rennes, France.
[2]INRA, UMR598, Animal Genetics, 35000 Rennes, France.
[3]Agrocampus Ouest, Applied Mathematics Department, 35000 Rennes, France.
Corresponding author: yuna.blum@rennes.inra.fr

Microarray technology allows the simultaneous analysis of thousands of genes within a single experiment. Classical approaches to analyze transcriptomic data ignore the gene dependence structure. This leads to correlation among test statistics which affects a strong control of the false discovery proportion.

We focus our study on a method called FAMT (Friguet et al, 2009) which captures the components of expression heterogeneity into factors. The relevance of factor modeling is first shown on illustrative gene expression data sets in simple situations of heterogeneity. We also use a real expression data set, primarily generated to map QTL for abdominal fatness in chickens (Le Mignon et al, 2009). The genelist found through FAMT was more related to the fatness trait that the one found by the classical way. Indeed, a PCA generated with the FAMT genelist discriminates much more fat and lean chickens. FAMT provides also functional information about a QTL region through a gene related to the fatness trait and controlled by this region (DHCR7) not observed by a classical approach. Then we interpret the independent factors extracted from this biological data set using known information about both experimental design and genes. We show that some factors may have different and complex origins, which can be related to particular metabolisms.

As we extract biological information from what was before simply considered as statistical noise, analyzing heterogeneity in gene expression yields a new point of view on transcriptomic data (Blum et al, submitted).

P2004 GenotypeChecker: an interactive tool for checking the inheritance consistency of genotyped pedigrees

Trevor Paterson and Andy Law.
The Roslin Institute, Royal (Dick) School of Veterinary Studies,
The University of Edinburgh, Edinburgh EH25 9PS, UK.
Corresponding author: trevor.paterson@roslin.ed.ac.uk

The raw data for performing genetic analyses consists of pedigree information about related individuals together with genotype data for any number of detectable polymorphic markers. The pedigrees analyzed may be large in size and complex in structure. Furthermore, the emergence of techniques such as microarray SNP-genotyping has led to an explosion in the scale of data, where 1000s of markers may be genotyped across the population.

Downstream processing of such data, for example to analyze linkage associations between markers, is highly sensitive to errors in the data. Errors in the recorded pedigree structure or incorrect genotypes generate inconsistencies with the rules of Mendelian inheritance, which invalidate or skew genetic analysis algorithms.

Identifying and cleaning data errors prior to downstream processing is complex for such large datasets and, because genotype or identification errors have consequences on the apparent inheritance pattern of relatives, it is difficult to unambiguously identify exact source errors.

'GenotypeChecker' allows users to load and analyze large pedigree/genotype datasets and check the data for inheritance inconsistencies. The tool combines a genotype checking algorithm which applies the rules of Mendelian inheritance across all markers for the pedigree, with an interactive exploratory interface that visualizes the inheritance of markers in the context of the pedigree structure. The data are presented to the user in colour-coded tables which can be sorted and filtered to explore the reported errors. The user can 'mask' individual suspect genotypes prior to rechecking inheritance, providing verification of erroneous datapoints that should be cleaned from the data.

Funded by BBSRC (BBS/B/05478/2).