Arch. Tierz., Dummerstorf **49** (2006) Special Issue, 110-115

Wageningen University and Research Centre, Animal Sciences Group, Division of Animal Resource Development, ID-Lelystad, PO Box 62, 8200 AN Lelystad, The Netherlands
*: Present address: ETH Statistical Animal Genetics Group, ETH Zentrum, Universitätsstrasse 65 (UNS D7), CH-8092 Zürich, Switzerland

MARINUS F.W. TE PAS, MARCO H. POOL, INA HULSEGGE and LUC L.G. JANSS*

Analysis of the differential transcriptome expression profiles during prenatal muscle tissue development in pigs

Abstract
In this contribution two microarray experiments are reviewed aiming to describe (1) the differences in the expression profiles of Duroc and Pietrain pigs during prenatal muscle tissue development, and (2) The changes in the expression profiles of genes related to myogenesis in Duroc pigs. Furthermore, we describe bioinformatics and pathway analyses methods to extract biological meaningful knowledge.

Key Words: Pig breeds, Myogenesis, Microarray, Bioinformatics, Pathway analysis

Zusammenfassung
Titel der Arbeit: **Analyse der differentiellen Transkript-Expressionsprofile während der pränatalen Muskelentwicklung beim Schwein**
Der Beitrag erläutert zwei Microarray Experimente die darauf abzielen, (1) die Unterschiede in den Expressionsprofilen von Duroc- und Pietrainschweinen während der prenatalen Muskelentwicklung aufzuzeigen und (2) die Änderungen der Expressionsprofile der Gene zu beschreiben, die in der Myogenese bei Durocschweinen relevant sind. Außerdem beschreiben wir Bioinformatik- und Pathway-Analyse-Methoden, um biologisch sinnvolle Erkenntnisse abzuleiten.

Schlüsselwörter: Schweinezucht, Myogenesis, Microarray, Bioinformatik, Pathway-Analyse

Introduction
Mammalian myogenesis, the formation of new multinucleated muscle fibres from mononucleated precursor cells called myoblasts, is an exclusive prenatal process determining muscle characteristics such as fibre numbers, which may be related to muscle strength and function (REHFELDT et al., 2000). Muscle fibre formation takes place in two waves, the primary and secondary muscle fibre formation (WIGMORE and EVANS, 2002). Each wave consists of proliferation of myoblasts and fusion to form new muscle fibres. While primary muscle fibres form de novo, secondary myofibres form using the primary fibres as a template.
Myogenesis is under complex genetic regulation. The Muscle Regulatory Factors (MRF) gene family are known to be transcription factors activating muscle-specific genes during different stages of myogenesis (OLSON, 1990; WEINTRAUB et al., 1991). The expression of the MRF genes is under tight temporal and spatial regulation, and numerous factors affecting MRF expression levels are known. A network of genes affects the expression patterns of the MRF genes (OLSON, 1993; RAWLS and OLSON, 1997; CAPDEVILA and JOHNSON, 2000; DOBOSY and SELKER, 2001; KITZMANN and FERNANDEZ, 2001; LEE et al., 2001; ZHU et al., 2001). By doing so they affect muscle and body growth potential.
Pig breeding has mainly focused during the past decades on improving growth rate and muscularity (MERKS, 2000). Pig breeds differ in muscle traits such as muscularity, muscle fibre type, colour, etc. For example, Duroc are slow growing pigs with a relatively high intramuscular fat content (SELLIER, 1998) and relatively red muscle

fibre types. Pietrain pigs are faster growing pigs with relatively low intramuscular fat content (JONES, 1998; SELLIER, 1998) and whiter muscle fibre types. Also overall fatness of Duroc pigs is greater than Pietrain pigs. These two breeds are considered to represent extremes of modern western pig breeds. It can be expected that differences in myogenesis are a major underlying mechanism for the observed phenotypes.
Using microarray technology we studied the porcine expression of genes known to affect myogenesis in laboratory animals and *in vitro* model systems. Microarray technology can simultaneously measure the differential expression of a large number of genes in a given tissue and may identify the genes involved in different phenotypes. Typically, microarray experiments produce long lists of genes that are differentially expressed between two different situations. This does not necessarily shed light on the underlying complex genomic regulation that creates the different phenotypes. Here we discuss several of our experiments to elucidate the genetic background of the myogenesis in pigs and the differences between pig breeds (TE PAS et al., 2005a,b; CAGNAZZO et al., 2006), and we extend these studies using bioinformatics.

Materials and Methods

Microarrays and analyses:
For details of the experiments see TE PAS et al. (2005a,b) and CAGNAZZO et al. (2006).

Pathway analysis
For pathway analysis we used the KEGG (Kyoto Encyclopaedia of Genes and Genomes) and BioCarta data bases (resp. http://www.genome.ad.jp/kegg/, and http://www.biocarta.com/). The databases were searched with lists of differentially expressed genes using either home made software (http://www.do.asg.wur.nl/research/researchprojects2.asp?projectnr=105) or GoMiner (http://discover.nci.nih.gov/gominer/). Furthermore, the Spotfire software package (http://www.spotfire.com/) was used to investigate for putative pathways that are presently not covered by either of the two databases.

Results

Comparison of the Myogenesis-associated transcriptomes of Duroc and Pietrain (TE PAS et al., 2005; CAGNAZZO et al., 2006)
Genes were grouped into three major groups: myogenesis, energy metabolism, and muscle structural genes. Results were analysed for (1) up / down regulation - i.e. the log ratio between the expression level in Duroc and the expression level in Pietrain (M-value), and (2) for general expression level (average of log intensities or A-value).
Energy metabolism. A major difference between the energy metabolism in Duroc and Pietrain embryos and foetuses was observed at all gestational ages. The energy metabolism in the Pietrain is at a higher level than in the Duroc except 35 d of gestation where a reversed situation was found.
The expression profile of the genes of the fatty acid metabolism indicate that fatty acid metabolism is at a higher level in early Duroc embryos (d 14 - 49 of gestational age) compared to Pietrain embryos while the reverse situation is found in older foetuses from d 63 of gestation and onwards suggesting that the observed overall higher fatness in Duroc compared to Pietrain already develops in the early embryo.

Myogenesis. The transcriptome expression profiles of myogenesis related genes suggests that myogenesis starts up earlier in Duroc than in Pietrain. From 49d and onwards with the exception of the differentiation-inhibiting group of genes Pietrain shows increased myogenesis. The results suggest that early primary myogenesis more relates to Duroc muscle fibre formation while Pietrain muscles are more formed during secondary muscle fibre formation. The expression profiles of the *muscle structural genes* support the results of the myogenesis.

The Duroc-specific Myogenesis-associated transcriptome profiles (TE PAS et al., 2005a, b)

The genes involved in muscle fiber formation, i.e. differentiation-stimulating, differentiation-inhibiting, and muscle fiber structural genes, show a peak expression around day 35. The genes regulating myoblast proliferation show lower peak levels a few days earlier. The gene-activation profiles coincide in general with these profiles. Together the results suggest that the switch from myoblast proliferation to myoblast fusion (differentiation) is regulated by a decreasing number of expressed proliferation regulating genes and an increasing number of expressed differentiation regulating genes.

Furthermore, expression of glycolysis metabolism genes is at a nadir at the two period's central in differentiation: around days 35 and 49-63. ATP metabolism follows that profile later in time while oxidative phosphorylation has less variable expression. Together these results suggest that energy metabolism is coupled to myoblast proliferation and differentiation, but a possible causal relationship remains unclear.

Pathway analysis

Genes with known pathway information

While these results show interesting details associated with the regulation of myogenesis, the biological mechanisms that are active during those processes remain poorly understood. It will be necessary to take the lists of genes that are differentially regulated and compare them with physiological data to find the biological active components of these processes. That is where we started to analyse the active pathways of those genes. Physiological pathway information was extracted from the KEGG and BioCarta databases. The genes that were active in each represented pathway were grouped together and the myogenesis-associated transcriptome profiles in Duroc were compared within each relevant pathway. Over 20 different pathways were found to be active in myogenesis. The number of genes within each pathway varied from only two genes to more than 20 genes. Individual gene expression profiles were recorded over developmental age as the M-values for each two successive prenatal ages. Figure 1 shows two examples. In Figure 1A the TGF-β pathway represented by 12 genes shows that all genes in the pathway have a peak level at 35-49 days of age followed by a decrease in expression to a nadir at days 63-77 and an increase towards the end of the profile at day 91 of prenatal development. The beta-Catenin pathway in Figure 1B represented by 20 genes does not show such an ordered profile. Several genes show peak levels where others show nadirs, and no general profile can be observed. If this occurs we ordered the genes in profiles such that the genes within a profile have a similar expression profile. In the beta-Catenin pathway at least two profiles can be described. One profile with genes showing a peak value around day 40, a nadir around day 60 and a peak at day 91, a second profile shows

exactly the opposite profile. Additionally, a third profile could be assigned with genes showing no differential expression from day 14 to day 50, a nadir around day 60 and a peak level at the end of the profile. This suggests that this pathway could be seen as three independent pathways (related to the expression profiles) each independently regulated.

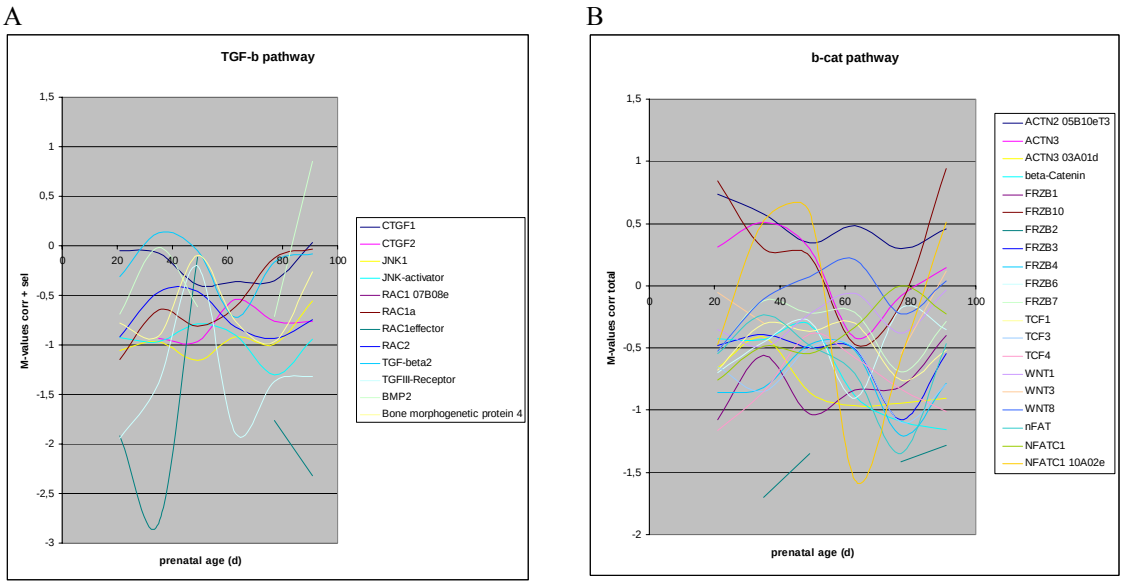

Fig. 1: Expression profiles of genes in two pathways influencing myogenesis: The TGF-β (A) and the beta-Catenin (B) pathways (Expressionsprofile von Genen zweier Reaktionswege der Myogenese: TGF-β (A) und the beta-Catenin (B) Pathways)

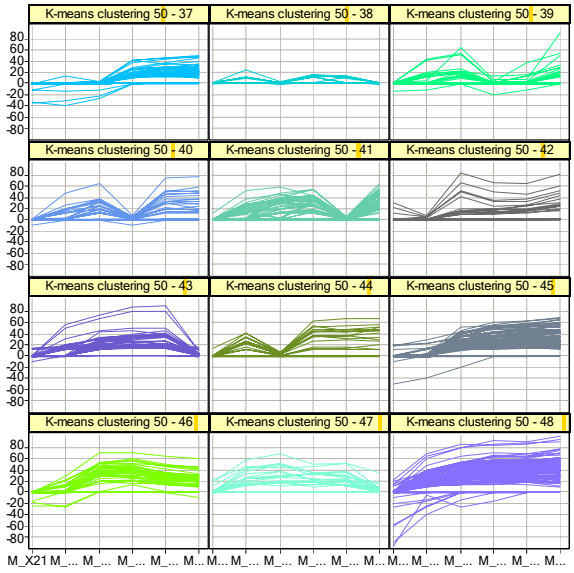

Fig. 2: Cluster analysis using K-means clustering (Klusteranalyse mit dem K-means Algorithmus)

Genes without known pathway information

Many genes on the microarray have no information in pathway databases, either because no information is available, or because the identity of the spot remains unknown. To improve our knowledge about the process it would be interesting to add pathway information of these genes to the myogenesis regulating pathways. Cluster analysis using the K-means clustering method can cluster genes with similar

expression profiles together. It can be assumed that genes with similar expression profiles have similar regulation of expression and can therefore be in common pathways or in pathways that act together. Figure 2 shows an example of part of a clustering analysis. It can be seen that the genes within a cluster have very similar expression profiles while major differences exist between the expression profiles of the genes in different clusters. Each cluster consists of spots of genes with known pathway information and genes without such information. It can be assumed that genes that cluster together act either in the same pathway or act in pathways that function together.

Discussion

With the onset of the "omics" sciences biology moved from studying isolated steps of a process towards integration of the study of live as a whole. However, the information content of the "omics" experiments is huge and it is a major task to extract biological meaningful knowledge. With the integration of the diverse "omics" sciences (transcriptomics, proteomics, metabolomics) this task will only become more complex. In the experiments that we have reported on myogenesis we only studied transcriptomic data, and we are still analysing the results in order to understand the biology behind the data. In this paper we shortly reviewed the previously published experiments and we indicated possible routes that we are working on to understand the biology of the regulation of myogenesis. Pathway information tells us how genes act together in a single or multiple pathways to perform myogenesis. Pathway information may tell us how the expression of all these genes is regulated in a large system. Understanding pathways and interactions between pathways may give us understanding on the functioning of the genome. These analyses are still going on but the first results are promising: we can now see that some pathways are most probably regulated as a single unit while others seem to have more complex regulation.

Although physiology has uncovered many pathways still many other pathways are only poorly known, whilst others remain unknown. Therefore, we are trying to add more genes to known and unknown pathways. Cluster analysis groups genes with similar expression profiles suggestive for similar regulation. Although no proof is given that genes within a cluster belong to the same pathway such results may be the starting point for further research. Adding more pathways of life to our knowledge will help us understand how life functions.

Acknowledgement

This work was financially supported by an EU grant under contract no. QLK5-CT-20000-01363 with contra finances from the Director of ID-Lelystad. The pathway analyses project was financially supported by an internal grant of Wageningen University and Research Centre.

References

CAGNAZZO, M.; TE PAS, M.F.W.; PRIEM, J.; DE WIT, A.; POOL, M.H.; DAVOLI, R.; RUSSO, V.:
Comparison of prenatal muscle tissue expression profiles of two pig breeds differing in muscle characteristics. J. Anim. Sci. (2006), in press.

CAPDEVILLA, J.; JOHNSON, R.L.:
Hedgehog signaling in vertebrate and invertebrate limb patterning. Cell. Mol. Life Sci. **57** (2000), 1682-1694

DOBOSY, J.R.; SELKER, E.U.:
Emerging connections between DNA methylation and histone acetylation. Cell. Mol. Life Sci. **58** (2001), 721-727

JONES, G.F.:
Genetic aspects of domestication, Common breeds and their origin. In: The Genetics of the Pig (Eds M.F. ROTHSCHILD and A. RUVINSKY), CAB International, Oxon, U.K. (1998), 640 p.

KITZMANN, M.; FERNANDEZ, A.:
Crosstalk between cell cycle regulators and the myopgenic factor MyoD in skeletal myoblasts. Cell. Mol. Life Sci. **58** (2001), 571-579

LEE, J.W.; LEE, Y.C.; NA, S.-Y.; JUNG, D.-J.U; LEE, S.-K.:
Transcriptional coregulators of the nuclear receptor superfamily: coactovators and corepressors. Cell. Mol. Life Sci. **58** (2001), 289-297

MERKS, J.W.M.:
One century of genetic changes in pigs and the future needs. In: The challenge of genetic change in animal production (eds. W.G. HILL, S.C. BISHOP, B. MCGUIRK, J.C. MCKAY, G. SIMM, and A.J. Webb). Occasional publication no. 27. Brit. Soc. Anim. Sci., Edinburgh (2000). pp. 8-19

OLSON, E.N.:
Myo D family: A paradigm for development? Genes Dev. **4** (1990), 1454-1461

OLSON, E.N.
Signal transduction pathways that regulate skeletal muscle gene expression. Mol. Endocrinol. **7** (1993), 1369-1378

RAWLS, A.; OLSON, E.N.:
MyoD meets its maker. Cell **89** (1997), 5-8

REHFELDT, C.; FIEDLER, I.; DIETL, G.; ENDER, K.:
Myogenesis and postnatal skeletal muscle cell growth as influenced by selection. Livest. Prod. Sci. **66** (2000), 177-188

SELLIER, P.:
Genetics of meat and carcass traits. In: The Genetics of the Pig (Eds M.F. ROTHSCHILD and A. RUVINSKY), CAB International, Oxon, U.K. (1998), pp. 463-510

TE PAS, M.F.W.; CAGNAZZO, M.; DE WIT, A.A.C.; PRIEM, J.; POOL, M.H.; DAVOLI, R.:
Muscle transcriptomes of Duroc and Pietrain pig breeds during prenatal formation of skeletal muscle tissue using microarray technology. Arch. Tierz., Dummerstorf **48** (2005a) Special Issue, 141-147

TE PAS, M.F.W.; DE WIT, A.A.C.; PRIEM, J.; CAGNAZZO, M.; DAVOLI, R.; RUSSO, V.; POOL, M.H.:
Transcriptome expression profiles in prenatal pigs in relation to myogenesis. J. Muscle Res. Cell Motil. (2005b), DOI 10.1007/s10974-005-7004-6

WEINTRAUB, H., DAVIS, R.; TAPSCOTT, S.; THAYER, M.; KRAUSE, M.; BENEZRA, R.; BLACKWELL, T.K.; TURNER, D.; RUPP, R.; HOLLENBERG, S.; ZHUANG, Y.; LASSAR, A.:
The myoD gene family: Nodal point during specification of the muscle cell lineage. Science **251** (1991), 761-766

WIGMORE, P.M.C.; EVANS, D.J.:
Molecular and cellular mechanisms involved in the generation of fiber diversity during myogenesis. Int. Rev. Cytol. **216** (2002), 175-232

ZHU, W.; FOEHR, M.; JAYNES, J.B.; HANES, S.D.:
Drosophila SAP18, a member of the Sin3/Rpd3 histone deacetylase complex, interacts with bicoid and inhibits its activity. Dev. Genes Evol. **211** (2001), 109-117

Corresponding Author
MARINUS F.W. TE PAS
Wageningen University and Research Centre, Animal Sciences Group,
Division of Animal Resource Development,
ID-Lelystad, PO Box 62, 8200 AN Lelystad, The Netherlands