Porcine imprinted genes detected by high-throughput cDNA-sequencing (RNA-seq)

Ole Madsen[*], Richard P.M.A. Crooijmans[*], Laurie A. Rund[#], Lawrence B. Schook[#], Martien A.M. Groenen[*]
[*]Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands
[#]Department of Animal Sciences, University of Illinois, Urbana, Illinois, USA
e-mail: ole.madsen@wur.nl

Introduction

- Genomic imprinting is an epigenetic phenomenon, where the level of expression of alleles depends on their parental origin.
- Studies in other mammals for imprinted genes are relative sparse and have generally been limited to either an *ad hoc* single gene analysis or analyses on "unnaturally" systems such as uniparental embryos.
- Some imprinted genes are imprinted in all tissues, whereas others are tissue or developmental stage specific imprinted.
- In eutherian mammals ~100 genes have experimentally been shown to be imprinted, mainly by studies in humans and mice.
- Comparative studies, mainly between human, mouse and marsupials, indicate that a relative large fraction of the imprinted genes are species-specific, suggesting that the evolution of genomic imprinting is a dynamic and ongoing process which is still poorly understood in details.
- A systematic investigation of imprinted genes in other placental mammals than mouse and human is therefore essential to achieve a detailed understanding of the evolution of imprinted genes in relation to their key role in brain development, regulation of growth and reproduction.

Aim

- The long term aim of this project is to detect the whole range of imprinted genes (the imprintome) in embryonic development of pigs.
  - Aim of this study is to make the first steps in developing the technical and analytic framework needed to detect the imprintome by means of high-throughput DNA and RNA (RNA-seq) sequencing.

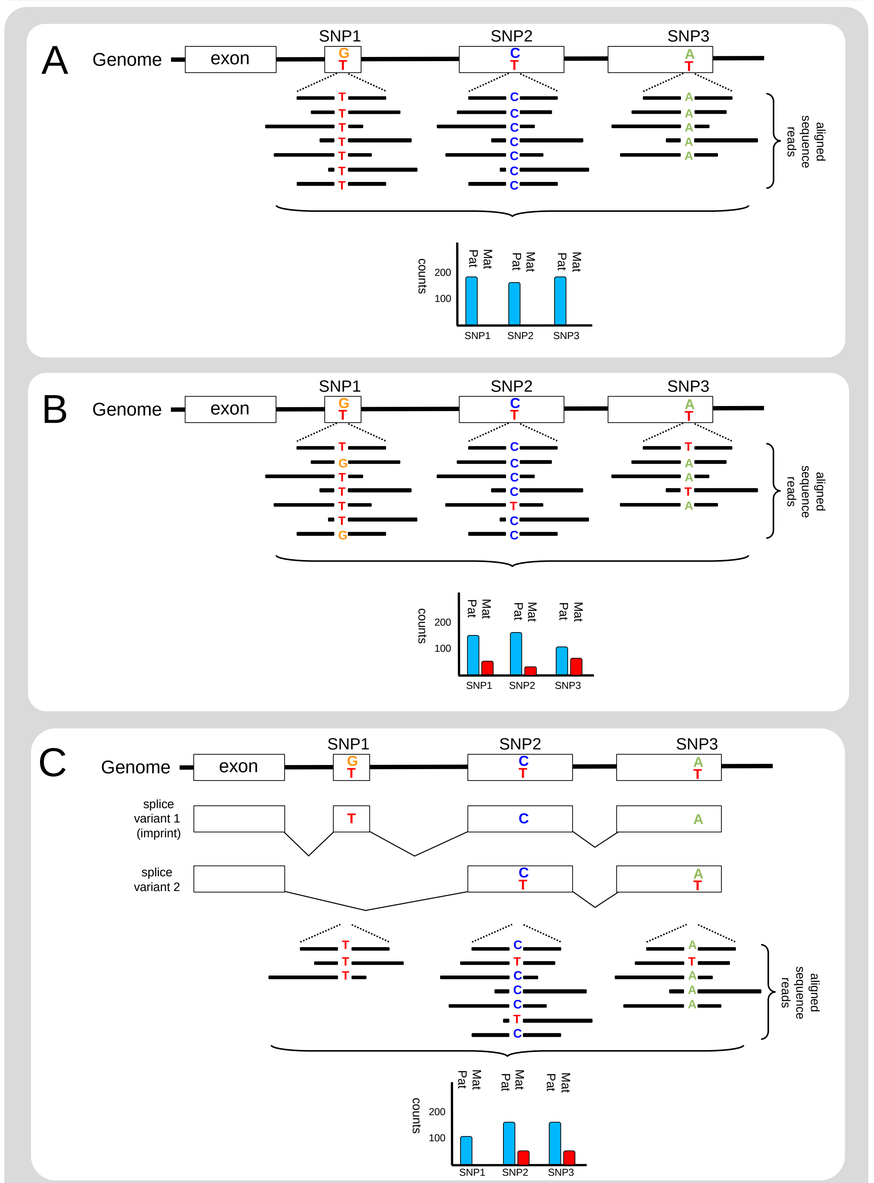
Material

- Illumina GA2 paired-end sequences of cDNA and genomic DNA of clones of the sequenced pig ('TJ Tabasco'):
  - ~23 million reads from placenta cDNA
  - (~120 million reads from a normalized cDNA library from a mix of 10 fetal tissues)
  - 24 x coverage of the pig genome for SNP discovery[1]

Method

- TopHat[2] (v.1.0.13 (BETA) with Bowtie v.0.10.5 implemented) was used to align cDNA reads to the pig reference genome (v9.56)
- Samtools (v0.1.7a) was used to score SNPs on the transcriptome alignments
- SOAP2 was used to align genomic reads to pig reference genome (v9.56) and SOAPsnp was used to score SNPs from aligned reads[1]

The principle



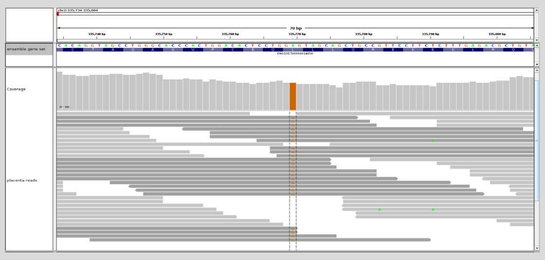
- The principle of detecting mono-allelic/imprinted gene expression by comparing a single individuals genomic variation with the variation in allelic expression in the same individuals transcriptome. Note that imprinting can only be detected if the genomic variation of both parents are available.

A) Complete mono-allelic or imprinted expression
B) Differential expressed or partly imprinted genes
C) Splice variant mono-allelic or imprinted expressed

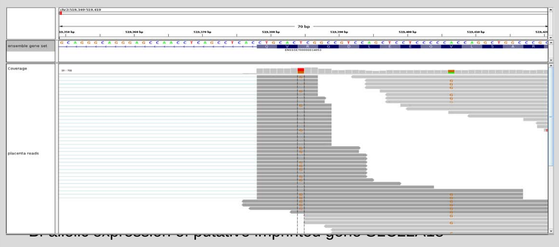
Results

- ?? genes known to be imprinted in mammals were investigated for mono-allelic expression by comparing the genomic variation in the sequenced pig ('TJ Tabasco') to the variation in allelic expression in the transcriptome of a clone of TJ Tabasco.
- ?? of these putative imprinted genes were missing in the current pig genome build, ?? were not expressed in the tested tissue, in ?? genes SNPs were lacking. Of the ?? genes which could be tested, ?? showed mono-allelic expression, ?? skewed allelic expression and ?? bi-allelic expression.

Here examples imprint, not imprint (imprint genes)



Mono-allelic expression of putative imprinted gene MEST



here table conservation

	Human	Mouse	Cow	Sheep	Pig	Marsupials	Monotreme
IGF2	+	+	+	+	+	+	-
IGF2R	-/+	+	+	+	+	+	-
NNAT	+	+	+	?	+	?	?
DLK1	+	+	?	+	+	-	?
ZIM2	+	+	-	?	?	?	?
GATM	-	+	?	-	?	?	?
UBE3A	+	+	?	?	?	-	-
BEGAIN	?	?	?	+	?	?	?

Conclusion

- Comparing of the genomic variation to the variation in allelic expression seems to be a valuable approach in detecting mono-allelic expressed genes as well as imprinted genes.
- The current preliminary results support the idea that the evolution of imprinted genes is a constantly ongoing process

Future plans

- Development of bioinformatic pipelines for detecting imprinted genes from next-generation sequences.
- Whole genome sequencing of family trios (parents and offspring), plus transcriptome sequencing of different fetal tissues from offspring to detect the whole imprintome in porcine
- Comparative analyses of mammalian imprinted genes

Acknowledgement

[1]We are very grateful to Drs. Chang Yu and LI Yingrui, BGI (China) and Dr. Mikkel H. Schierup, University of Aarhus (Denmark) for making the porcine genomic SNP available before publication.

This project is funded by the Dutch Technology Foundation (STW)

References:
1 Denoeud F, Aury JM, Da Silva C, Noel B, Rogier O, Delledonne M, Morgante M, Valle G, Wincker P, Scarpelli C, Jaillon O, Artiguenave F. Annotating genomes with massive-scale RNA sequencing. Genome Biology 2008, 9:R175; doi:10.1186/gb-2008-9-12-r175
2 Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 2009 25:1105-11; doi:10.1093/bioinformatics/btp120