WAGENINGEN UNIVERSITY
WAGENINGEN UR

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Detection of porcine imprinted genes through high-throughput cDNA-sequencing (RNA-seq)

Ole Madsen[*], Djawad Radjabzadeh[*], 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.
- In placental mammals >100 genes have experimentally been shown to be imprinted, mainly by studies in humans and mice. Recent studies indicates that as many as ~1300 loci are associated with parental specific allelic expression in mouse brain[1,2]
- 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 "unnatural" systems such as uniparental embryos.
- Some imprinted genes are imprinted in all tissues, whereas imprinting of other genes can be highly spatial and temporal with e.g. tissue or developmental stage specific imprinting.
- Comparative studies, mainly between human, mouse and marsupials, indicate that a relatively large fraction of placental imprinted genes are species-specific, suggesting that the evolution of genomic imprinting is a dynamic and ongoing process which is still poorly understood.
- 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 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 sequencing.

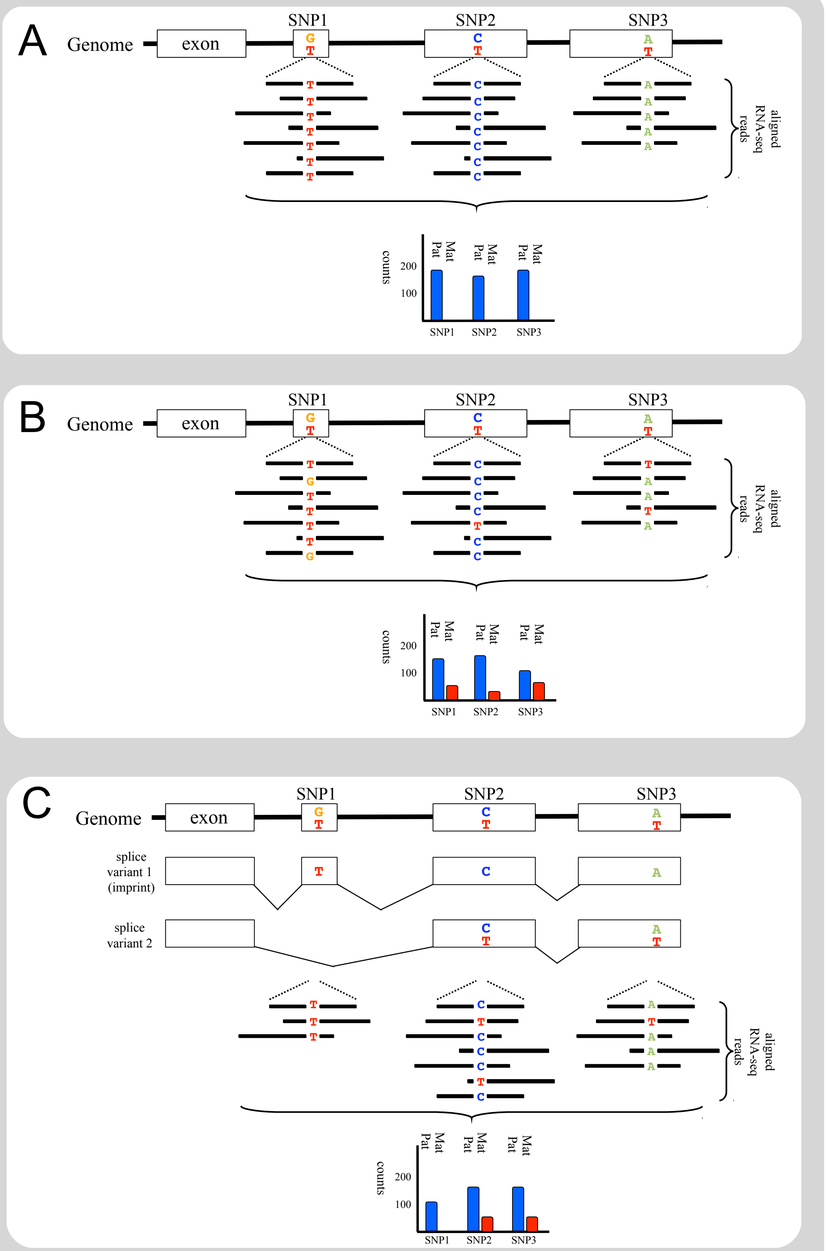
Material

- Illumina HighSeq paired-end sequencing of genomic DNA of two family trio's (father, mother and offspring at 8-12x coverage)
- Illumina HighSeq paired-end sequencing of total-RNA cDNA libraries from three offspring tissues (placenta, brain and liver at ~50x coverage for each sample)

Method

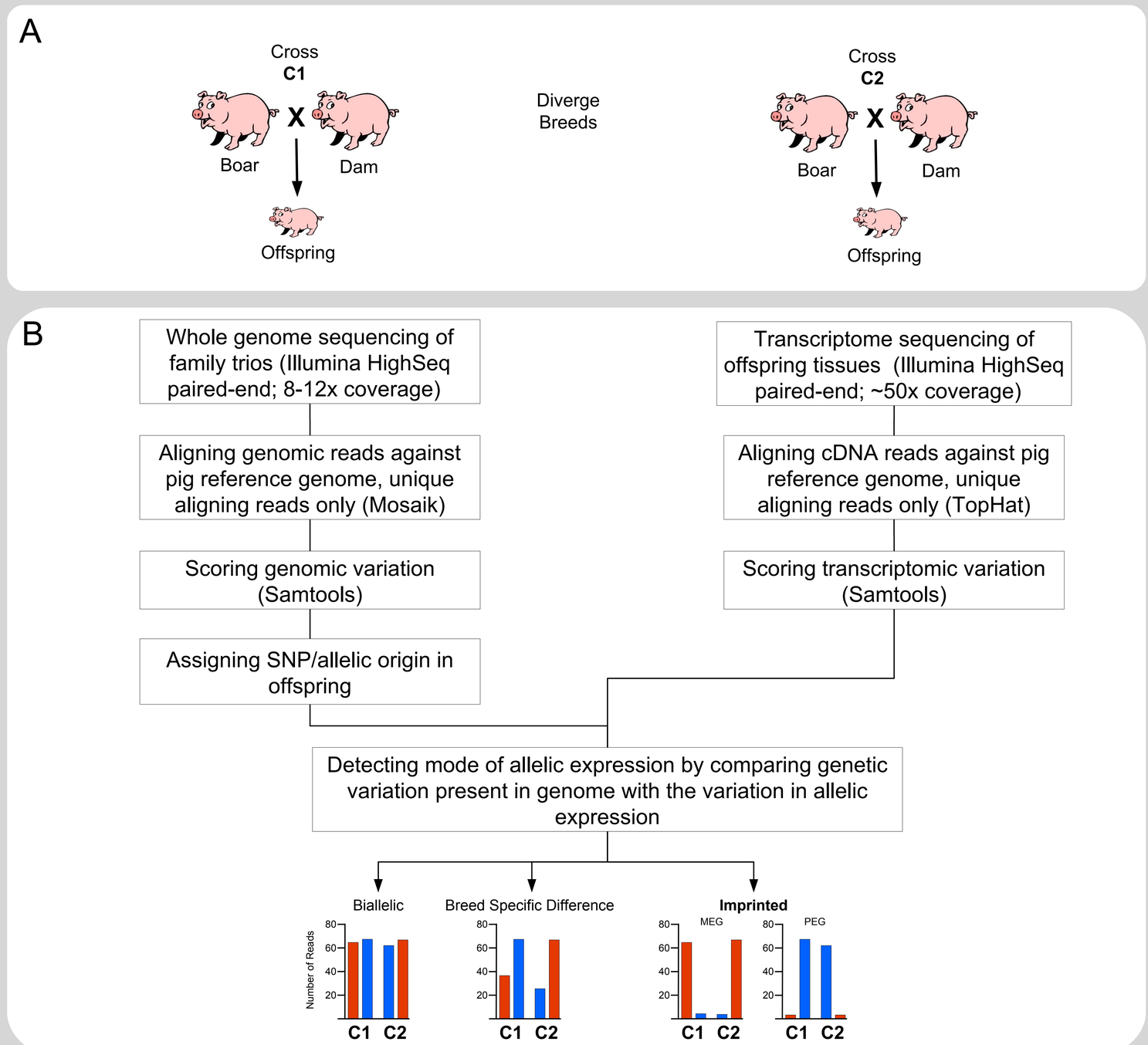
- TopHat[3] (with Bowtie implemented) will be used to align cDNA reads to the pig reference genome.
- Genomic reads will be aligned to the pig reference genome using Mosaik[4].
- SNPs from the transcriptome and genomic alignments will be scored using Samtools[5].

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

Setup Detecting Imprinted Loci



- A) Two family trios (C1 and C2) were created from cross of divergent pig breeds. Genomic DNA was isolated from trios and RNA from different tissues of offspring (day 60-65 of embryonic development)
- B) Flowchart for detecting the mode of allelic expression. MEG is Maternal Expressed Gene, PEG is Paternal Expressed Gene

References:
[1]Gregg et al. (2010) High-resolution analysis of parent-of-origin allelic expression in the mouse brain. Science **329**, 643-8.
[2]Gregg et al. (2010) Sex-specific parent-of-origin allelic expression in the mouse brain. Science **329**, 682-5.
[3]Trapnell C, Pachter L, Salzberg SL. (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25**, 1105-11.
[4]Li W and Strömberg M. (2010) MOSAIK. http://bioinformatics.bc.edu/marthlab/Mosaik
[5]Li H, et al. (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics **25**, 2078-9.

Acknowledgement

This project was supported by National Research Initiative Grant no. 2009-34480-19875 and 2006-35216-16668 from the USDA Cooperative State Research, Education; and by the Dutch Technology Foundation (STW). We would like to thank IPG and HyPor for supplying the two family trios.