Epigenetic mechanisms and their implications in animal breeding

Klaus Wimmers, Nares Trakooljul, Michael Oster, Eduard Murani, Siriluck Ponsuksili

Leibniz Institute for Farm Animal Biology (FBN), Germany
Institute for Genome Biology
The fate of cells

Zygote: a single cell with genetic information becomes a complex organisms with $10^{14}$ cells all containing the same genetic information.
At least two forms of information in the cell nuclei:

1. Genetic information: general instructions for the manufacture of all proteins – DNA sequence

2. Epigenetic information: additional instruction on how, when and where these information should be used – epigenetic marks
Genotype-Phenotype-Mapping

Environment + Genotype = Phenotype

Genome/Genotype/DNA

Phenotype
Epigenetic variation contributes to phenotypic variation; knowing it could improve the prediction of the phenotype

Epigenetic mechanisms link environment and genome: environment x genotype interactions
Use of the term epigenetics and its definition has changed throughout history.

**Conrad Waddington**, 1942 – study of epigenesis; how genotypes give rise to phenotypes in development


**Our current definition** (NIH Roadmap Epigenomics): Epigenetics is the study of mitotically (meiotically) heritable changes in gene expression that occur without changes in DNA sequence and of stable, long-term alterations of the transcriptional potential of a cell that are not necessarily heritable.
Biochemical reactions which are operating in Epigenetics

A. Modification at the DNA level
   cytosine methylation

B. Histone modification - the histone code
   histone acetylation, ~ methylation, ~ phosphorylation, ~ ubiquitination, ~ sumoylation
Histone modification

a

10-nm filament
Chromatin fibre
Higher-order folding
Nucleosome core particle

Unmodified histone H3
N-terminal ‘tail’
Globular domain
On ↔ Off
Modified histone H3

b

H3
N-ARTKQTAR
KSTGGKAPRKQLATKAARKSAP...

H4
N-SGRGKGGKGLGKGGAKRHRKVLRDNIQGIT...

Warsaw, Poland
Functional Annotation of Animal Genomes

FAANG

- Histone H3 lysine 4 trimethylation (H3K4me3), which correlates with promoters of active genes and transcription start sites;
- Histone H3 lysine 4 monomethylation (H3K4me1), which marks regulatory elements associated with enhancers and other distal elements, but is also enriched downstream of transcription start sites;
- Histone H3 lysine 27 trimethylation (H3K27me3), which marks genes that have been silenced through regional modification;
- Histone H3 lysine 27 acetylation (H3K27ac), which marks active regulatory elements, and may distinguish active enhancers and promoters from their inactive counterparts;
DNA methylation reactions

DNA-Methylation of CpG-di-nucleotides
→ causes stable gene inactivation
→ allows long-lasting gene expression control; imprinting
→ mechanism of acute gene regulation

DNMT – DNA Methyltransferases;
SAM – S-Adenosyl-Methionine

Strathdee and Brown 2002
DNA methylation reactions

Demethylase

DNA replication

De novo DNA methyltransferase (DNMT3a, 3b)

Maintenance DNA methyltransferase (DNMT1)

DNA methylation reactions
DNA-Methylation in Vertebrates

CpG-di-nucleotides ~1% of vertebrate genome

60-80% of all CpG are methylated (5% cytosines)

CpG-Islands: ≥200bp, ≥60% CpG
- promoter-associated; 50-60% of genes with CpG-Islands
- usually hypomethylated

CpG outside of CpG-islands strongly methylated
- Maintaining genomic stability
- long-term inactivation of repeats, retrotransposons,
- deamination of methCpG to TpG (C-T-Transition) !!!
Repression of Gene Expression by DNA-Methylation

1. Direct blocking of TFBS by methyl-group of CpG

2. Blocking of TFBS by methylcytosine binding proteins (MBP)

3. Recruitment of Histone Deacetylases by MBP leads to deacetylation of core histones → change in chromatin structure to heterochromatin

Iguchi-Ariga and Schaffner, 1989; Tierney et al., 2000; Ballestar and Wolffe, 2001; Nan et al., 1998; Fuks et al., 2003; D’Alessio and Szyf, 2006
Epigenetic phenomena

- X chromosome inactivation
- Genomic imprinting
- Gene inactivation (specific genes, transposable elements, repeats…)
- Tissue specific expression
- Acute regulation of expression
- Centrometric heterochromatin, organisation of chromatin
- Cancer
X chromosome inactivation

XIC = X inactivation center
XIST = X inactive specific transcript
17kb noncoding RNA
stable expressed from inactive X
“paints” inactive X chromosome (cis)
its own activity is affected by DNA methylation
X chromosome inactivation

zygote

mitosis

till blastocyst

random X-activation

at gastrulation

mosaic
Epigenetic reprogramming

DNA methylation

paternal genome
maternal genome
offspring somatic genome

mod. from Blewitt M., course slides
Epigenetic reprogramming

- paternal genome
- maternal genome
- offspring somatic genome
- repeats
Epigenetic reprogramming
Imprinting

diploid embryos derived from either only paternal or only maternal pronuclei failed to survive

Imprinted genes:
Mouse 132
cattle 25
Human 79
pig 21
sheep 14

• Mouse and human share only 40 imprinted genes
• Imprinting control region, ICR
• Often located in clusters
• Often coding for embryonic development, metabolism, behavior; relatively few but large effects

QTL with parent-of-origin effects

Surani, Barton, Norris, 1984
Imprinting of IGF2

Setting up boundaries. When the CTCF protein binds to DNA, it blocks regulatory DNA downstream from interacting with the Igf2 gene and only the H19 gene is expressed. If methyl groups (black) prevent CTCF binding, Igf2 is active, but H19 is silenced.

Rand & Cedar, 2003
Imprinting of IGF2

A>G transition in intron 3, CpG-island
A = Q, does not bind a repressor (ZBED6)
G = q, binds repressor; $G^{\text{met}}$ does not bind repressor
Explains 30% variance in lean meat
→ inheritance of the A-allele from the sire 3-fold increase in IGF2,

mod. from Rand & Cedar, 2003

Jeon et al., 1999; de Koning et al., 2000
Nezer et al., 1999; Van Laere et al., 2003
Imprinting of IGF2

- Sire line IGF2 A/A
- Dam line/grandparent boar IGF2 G/G
- Hybrid sow IGF2 G\textsuperscript{pat}/G
  - Improved: prolificacy, longevity
- Slaughter pig IGF2 A\textsuperscript{pat}/G
  - Improved: lean meat content, uniformity
Other implications of epigenetics

Genome + Epigenome $\Rightarrow$ Phenotype

Additional source variation:
Knowledge may contribute to predict the phenotype

- biomarker $\Rightarrow$ management tool
- selective breeding $\Rightarrow$ refinement of estimates
  imprinted locus = functionally hemizygous
Missing heritability & missing causality

- If the epimutation is stable:
  - Likely to be in LD with SNP (no implication for genomic selection)
- If epimutation is unstable:
  - No good selection criterion,
  - Does not contribute much to missing heritability
  - Identity by descent does not imply identity in state

But if epimutation is causal, it will not be detected by DNA sequencing

Goddard & Whitelaw, 2014
Slatkin, 2009
Gonzales-Recio, 2012
Other implications of epigenetics

- Assisted reproductive technologies: for example LOS

- Fetal programming (nutritional programming (conditioning)):

  Epidemiological data and experiments in model and farm animals revealed that environmental effects during gestation impact the phenotype of offspring

  Thrifty Phenotype Hypothesis

  Epigenetic mechanisms as a molecular memory are involved
**Impact of gestation diets**

<table>
<thead>
<tr>
<th>before mating</th>
<th>gestation</th>
<th>lactation</th>
<th>postweaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>low protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high carbohydrate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high fat</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- high methylating micronutrient

Oster, 2012
Epigenetic transgenerational inheritance

Generation 1
Experimental feeding to mid gestation

Generation 2

Generation 3
Experimental feeding to end gestation
Foetal Programming

- Impact of gestation diets with high and low protein content on gene expression
- Involvement of epigenetic mechanisms/DNA methylation

<table>
<thead>
<tr>
<th>LP</th>
<th>maternal low protein diet, 6.5% CP</th>
<th>Cross fostering</th>
<th>Standard diet</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>maternal adequate protein diet, 12.1% CP</td>
<td>Lactation diet, Litter size: n=11</td>
<td>ad libitum</td>
</tr>
<tr>
<td>HP</td>
<td>maternal high protein diet, 30% CP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Experiment 1: 8 sows per diet; liver: analysis of 8 sib pairs per diet muscle: analysis of 3 sib pairs per diet

Experiment 2: offspring from 12 sows per diet were distributed to postnatal sampling points liver: analysis of 8 sib pairs per stage and diet muscle: analysis of 3 sib pairs per stage and diet
• hierarchical influence of tissue, ontogenetic stage, and diet on transcript levels
• Muscle appeared to be a less resistant to nutritional modulation than liver
• no gatekeeper pathways GENes were obvious
• Differential expression of DNMT1, DNMT3a and DNMT3b
• Differential DNA-methylation of PPARα, NR3C1, CYP2C34, NCAPG…
DNA Methylation

Methionine metabolism

<table>
<thead>
<tr>
<th>Nutrient (mg/kg)</th>
<th>CON</th>
<th>MET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zinc</td>
<td>21.8</td>
<td>149.0</td>
</tr>
<tr>
<td>B6</td>
<td>3</td>
<td>1180</td>
</tr>
<tr>
<td>B12</td>
<td>31</td>
<td>5930</td>
</tr>
<tr>
<td>Folate</td>
<td>3</td>
<td>92.2</td>
</tr>
<tr>
<td>Choline</td>
<td>500</td>
<td>2230</td>
</tr>
<tr>
<td>Methionine</td>
<td>2050</td>
<td>4700</td>
</tr>
</tbody>
</table>
Epigenetic transgenerational inheritance

F2 offspring groups differed with respect to backfat percentage ($P = 0.03$) in liver and muscle differential expression of lipid metabolism and metabolic pathway. A significant difference in DNA methylation at the IYD gene

Raunschweig et al. 2012
Methylating micronutrients supplementation

To study the effects of methylating nutrient enriched maternal diet on DNA methylation and transcriptome changes

To map differentially methylated regions associated with treatment factors developmental stage, breed and maternal diet

To explore biological significance of DNA methylation changes by integration of DNA methylation and transcription profiles

2 breeds (Pi, GL) × 2 diets × 2 stages (91 ppc, 150 dpn)
4 offspring/breed/diet/stage, n=32, liver

Next Generation Sequencing:
RNA→Transcriptome (RNA-Seq)
DNA→Methylome (RRBS)
Genome coverage of the RRBS library

Frequency of annotated features in the pig genome.

Percentage of genomic features covered by greater than 4 reads.

RRBS effectively target specific genomic regions including CpG islands and CpG-island shores (±1kb from CpG island) and CpG rich promoters.
Percentage methylation of cytosines in CpG, CHG or CHH context

- 7.16% $^\text{m} \text{CpG}$
- 9.48% CpG
- 27.47% CHG
- 55.64% CHH
- 0.05% $^\text{m} \text{CHG}$
- 0.24% $^\text{m} \text{CHH}$

43.02% of CpGs are methylated

H = A, T or C
% = median of 32 libraries
Clustering based on the DNA methylation profile
differentially-methylated CpG site 1.9-kb upstream of the E2F7 transcriptional start site compared between control and methyl-donor rich maternal diet groups in DL and Pi pigs at 91-dpc stage.
Genome-wide distribution of DMRs associated with stage, breed and diet

DMRs with 15% differential methylation and q-value < 0.05.

433, 2038, and 932 DMRs were identified to be associated with stage, breed, and diet, respectively.
Functional annotation of genes in DMRs

A) Breed-associated DMRs/Genes

- Wnt/β-catenin-sign.
- Epithelial-mesenchymal transition
B) Maternal-diet-associated DMRs/Genes

Epithelial-mesenchymal transition

Eicosanoid Signaling
Human Embryonic Stem Cell Pluripotency
Role of Oct4 in Mammalian Embryonic Stem Cell Pluripotency

Wnt/β-catenin-sign.

Basal Cell Carcinoma Signaling
Hepatic Fibrosis / Hepatic Stellate Cell Activation
Autophagy
Vitamin-C Transport
Role of NANOG in Mammalian Embryonic Stem Cell Pluripotency
NGF Signaling

Adipogenesis

Adipogenesis pathway
Differential expression due to maternal dietary treatment

Methionine metabolism: regulated in both breeds including DNMT1, DNMT3a, DNMT3b
Differential expression due to maternal dietary treatment

- 150dpn:
  - several lipid metabolism pathways

- 91dpn:
  - GADD45 signaling
  - Pyridoxal 5-phosphate salvage pathway
  - Folate polyglutamylation
  - IGF signaling
  - Wnt/β-catenin signaling

426 and 363 genes differentially expressed at 91dpc and 150dpn, respectively.
Functional annotation of DMRs: Serine-protein kinase ATM locus

ATM The PI3/PI4-kinase family; cell cycle checkpoint signaling pathways; DNA damage and genome stability

Chr9:40,842,163 - 40,846,955
Integration of RRBS and RNA-Seq

Significant correlation between RNA-Seq and RRBS

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Symbol</th>
<th>Strand</th>
<th>Chr</th>
<th>Start</th>
<th>End</th>
<th>Upstream (kb)</th>
<th>Mean correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSSSCG00000016981</td>
<td>CPEB4</td>
<td>-1</td>
<td>16</td>
<td>54813538</td>
<td>54876701</td>
<td>-4.88</td>
<td>-0.3486</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-6.59</td>
<td>-0.6066</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-7.18</td>
<td>-0.5185</td>
</tr>
<tr>
<td>ENSSSCG00000002648</td>
<td>CBFA2T3</td>
<td>+1</td>
<td>6</td>
<td>1022607</td>
<td>1044687</td>
<td>-66.32</td>
<td>-0.4726</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-66.92</td>
<td>-0.4569</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-67.47</td>
<td>-0.4833</td>
</tr>
<tr>
<td>ENSSSCG00000005849</td>
<td>TUBB4B</td>
<td>-1</td>
<td>1</td>
<td>3141</td>
<td>35326</td>
<td>-78.00</td>
<td>-0.4348</td>
</tr>
<tr>
<td>ENSSSCG00000020149</td>
<td>5_8S_rRNA</td>
<td>-6</td>
<td>6</td>
<td>872546</td>
<td>872698</td>
<td>-0.28</td>
<td>-0.3231</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.71</td>
<td>-0.3150</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.09</td>
<td>-0.3378</td>
</tr>
<tr>
<td>ENSSSCG00000005857</td>
<td>C9orf169</td>
<td>-1</td>
<td>1</td>
<td>314147913</td>
<td>314151525</td>
<td>-64.34</td>
<td>-0.3924</td>
</tr>
<tr>
<td>ENSSSCG00000005856</td>
<td>RNF224</td>
<td>-1</td>
<td>1</td>
<td>314145387</td>
<td>314145860</td>
<td>-70.05</td>
<td>-0.3889</td>
</tr>
<tr>
<td>ENSSSCG00000022101</td>
<td>PIGBRCA1</td>
<td>-1</td>
<td>12</td>
<td>20029801</td>
<td>20091791</td>
<td>-88.71</td>
<td>-0.3786</td>
</tr>
<tr>
<td>ENSSSCG00000006348</td>
<td>FCRLB</td>
<td>-4</td>
<td>4</td>
<td>96850069</td>
<td>96855596</td>
<td>-86.52</td>
<td>-0.3754</td>
</tr>
<tr>
<td>ENSSSCG00000006350</td>
<td>CD32</td>
<td>-4</td>
<td>4</td>
<td>96888986</td>
<td>96908203</td>
<td>-31.97</td>
<td>-0.3574</td>
</tr>
</tbody>
</table>

Significant correlation between RNA-Seq and RRBS.
Association between DNA methylation state and gene expression suggesting biological significance of the identified DMRs.

Chr17:66519911 - 66519922; distance to C7SZ TSS: -124 bp

Chr3:chr3:41636669 - 41636745; distance to IGFALS TSS: 479 bp
Summary

• Epigenetic marks determine the transcriptional potential of the cell
• They represent an additional level of information to explain the phenotype
• Epigenetic marks link environment and genome and represent the molecular equivalent of genotype × environmental interactions
• Knowing and understanding epigenetic marks will contribute to the refinement of estimates in selective breeding
• Transgenerational inheritance and the contribution to “missing heritability” are a matter of debate
• Epigenetic studies will provide biomarker for management and/or breeding
• It will be possible to promote targeted epigenetic marks to develop certain phenotype
Thanks to my co-authors and lab-team

SABRE (EU-FP6)
JP. Renard, INRA, FR
M. Braunschweig, Uni Bern, CH

FEPROeXPRESS
CC. Metges FBN
H. Sauerwein, Uni Bonn

Thank you for your attention!!!