Whole liver transcriptome analysis for the metabolic adaptation of dairy cows

Ngoc-Thuy Ha
Animal Breeding and Genetics Group
Department of Animal Sciences
Georg-August-University Goettingen, Germany

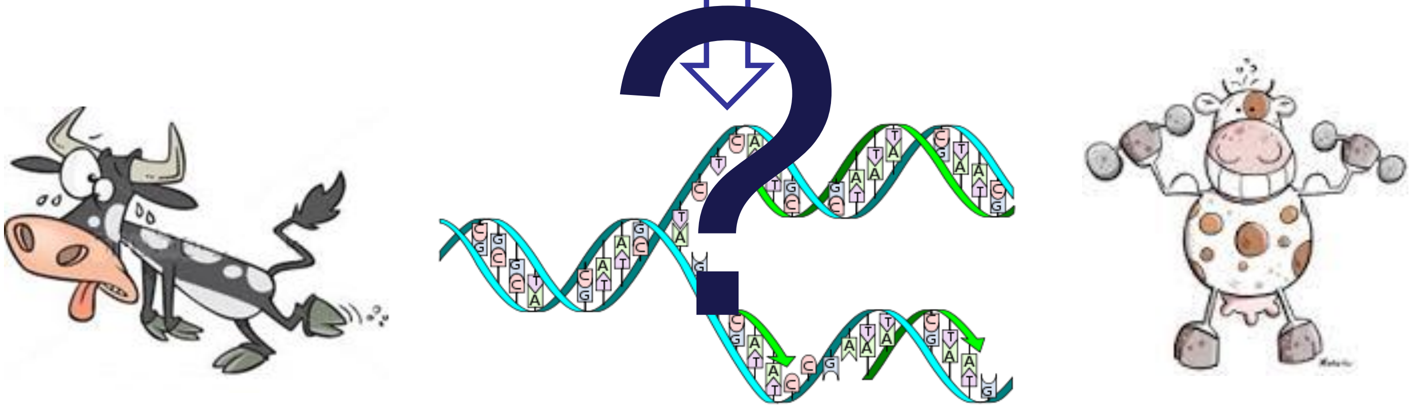
Motivation – ***Background***

- Selective breeding of high-yielding dairy cows
  - ⇒ up to 60 kg milk per day
- High energy demand can not be fully covered by food intake
  - ⇒ negative energy balance during their early lactation
- Mobilization of body fat, protein and mineral stores
  - ⇒ adaptation of the hepatic metabolism

Motivation – *Metabolic Adaptation*

Why does the success of adaptation differ substantially between cows - even under the same conditions and similar production levels?

Ingvartsen et al. (2003)
Drackley et al. (2005)
Graber et al. (2010)



Project OnlyRobust – *Overview*

Goal: Identification of candidate genes and pathways important for the metabolic adaptation

Study 1: Genomic Level:

„Gene-based mapping and pathway analysis of metabolic traits in dairy cows“ (Ha et al., 2015)

Study 2: Transcriptomic Level:

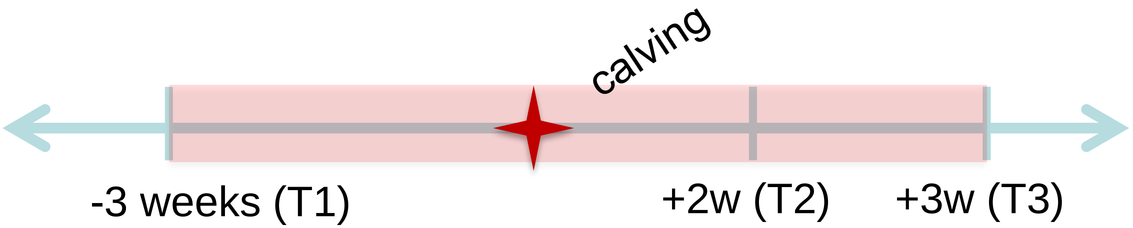
„Whole liver transcriptome analysis for the metabolic adaptation of dairy cows“ (in preparation)

Data – ***Experiment Design***

Identification of genes und pathways, important for metabolic adaptation during the transition period

Data:

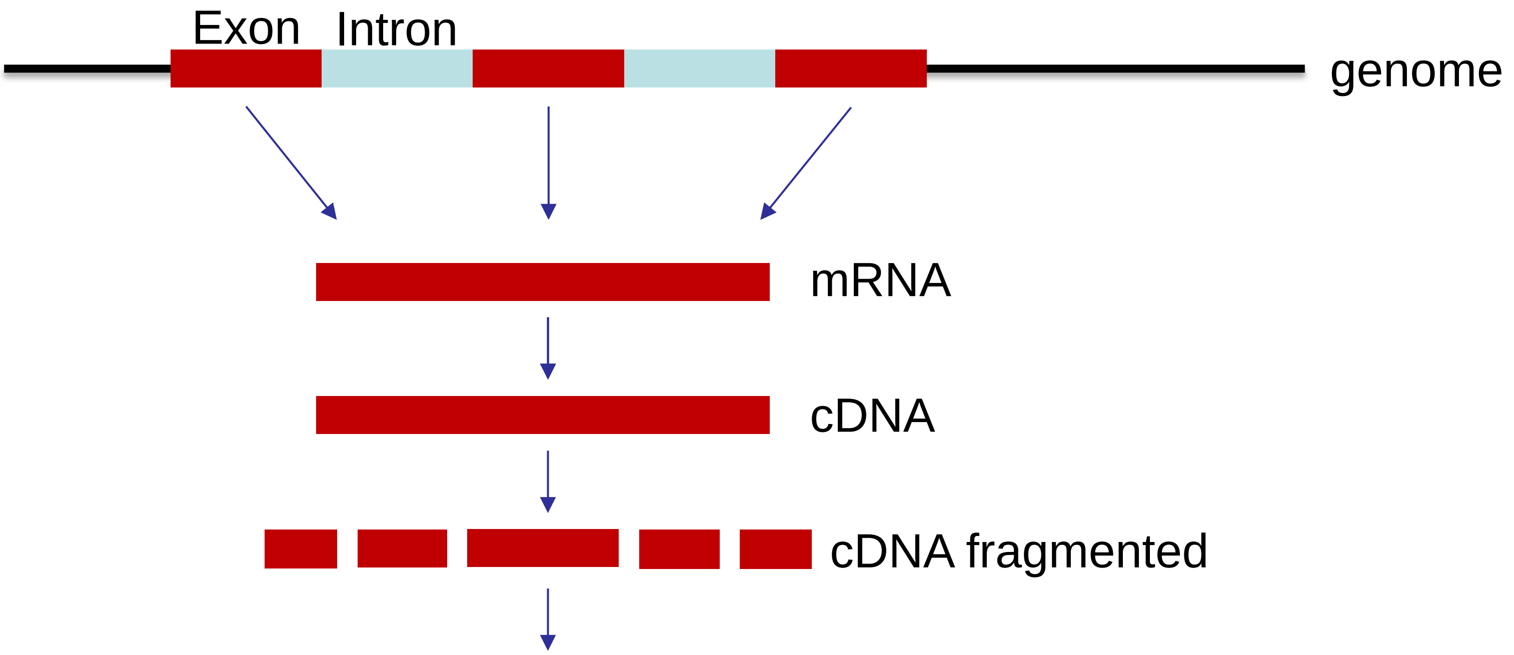
- *Liver samples of 6 dairy cows at different points of time:*



- *RNA-Sequencing*

Differential Gene Expression and Pathway Analysis:
T1 vs. T2
T1 vs. T3
T2 vs. T3

Data – *RNA-Seq*



Exon Intron

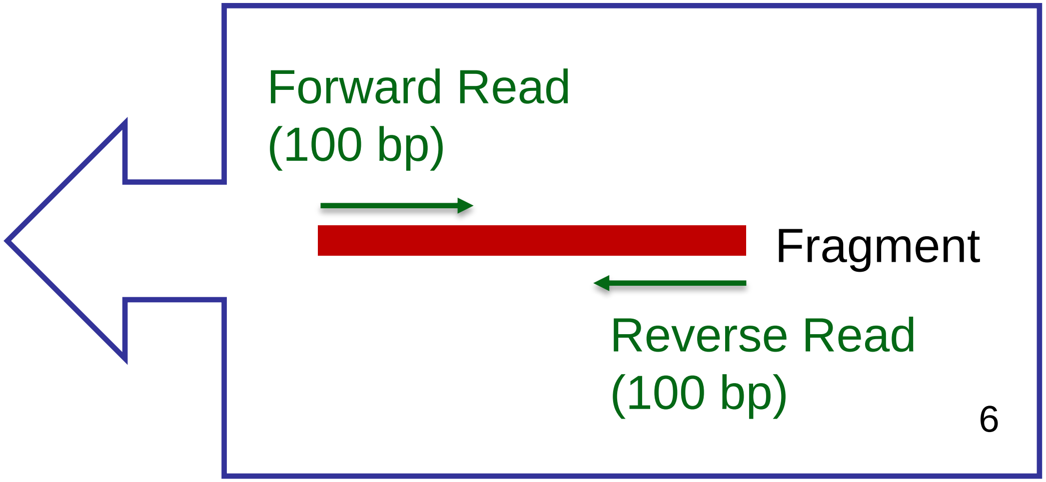
genome

mRNA

cDNA

cDNA fragmented

Illumina HiSeq 2000

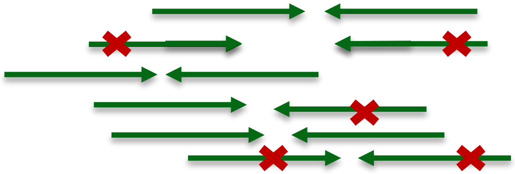


Forward Read
(100 bp)

Fragment

Reverse Read
(100 bp)

RNA-Seq – Pipeline



(Trimmomatic V0.33, Bolger et al., 2013)

(STAR V2.4,
Dobin et al., 2013)

Bos Taurus
(UMD 3.1, release 85)

2 Counts

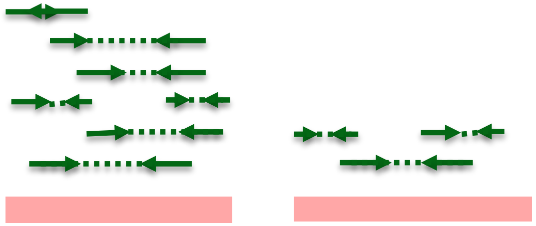
4 Counts

Count-Matrix (after mapping and counting of the reads):

	Sample 1	**Sample 2**	**Sample 3**	**…**
Gene 1	26‘432	20‘422	19‘543	…
Gene 2	0	24	3	…
Gene 3	245	432	332	…
…	…	…	…	…

Analysis of RNA-Seq

Differential Gene Expression Analysis (DGEA)

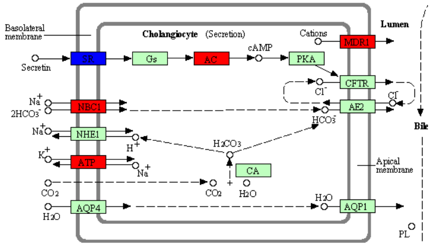


ante-partum

post-partum

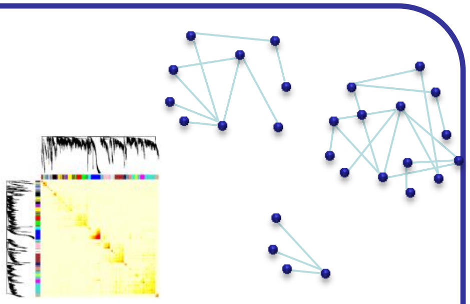
List of differentially expressed genes

Gene Set Enrichment Analysis (GSEA)



KEGG pathways

Weighted Gene Co-expression Network Analysis (WGCNA)



Methods – *DGEA and GSEA*

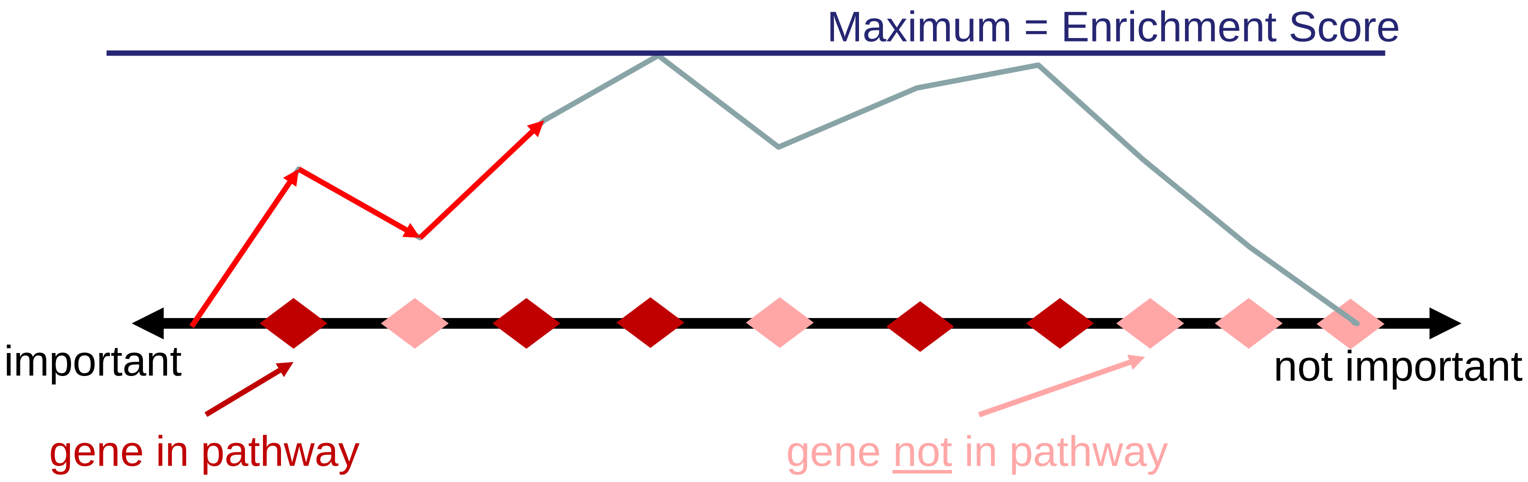
Differential Gene Expression Analysis (DGEA):

- *Generalized Linear Model using negative Binomial distribution („edgeR“, Robinson und Smyth, 2008)*

$$\log(Y) = \mu + \tau_1 \mathrm{cow}_1 + \cdots + \tau_6 \mathrm{cow}_6 + \beta \mathrm{condition} + \log N$$

- *Y = #counts per gene, N = #counts for all genes in a sample*

Gene Set Enrichment Analysis (GSEA, Subramanian et al., 2005):



Maximum = Enrichment Score

important

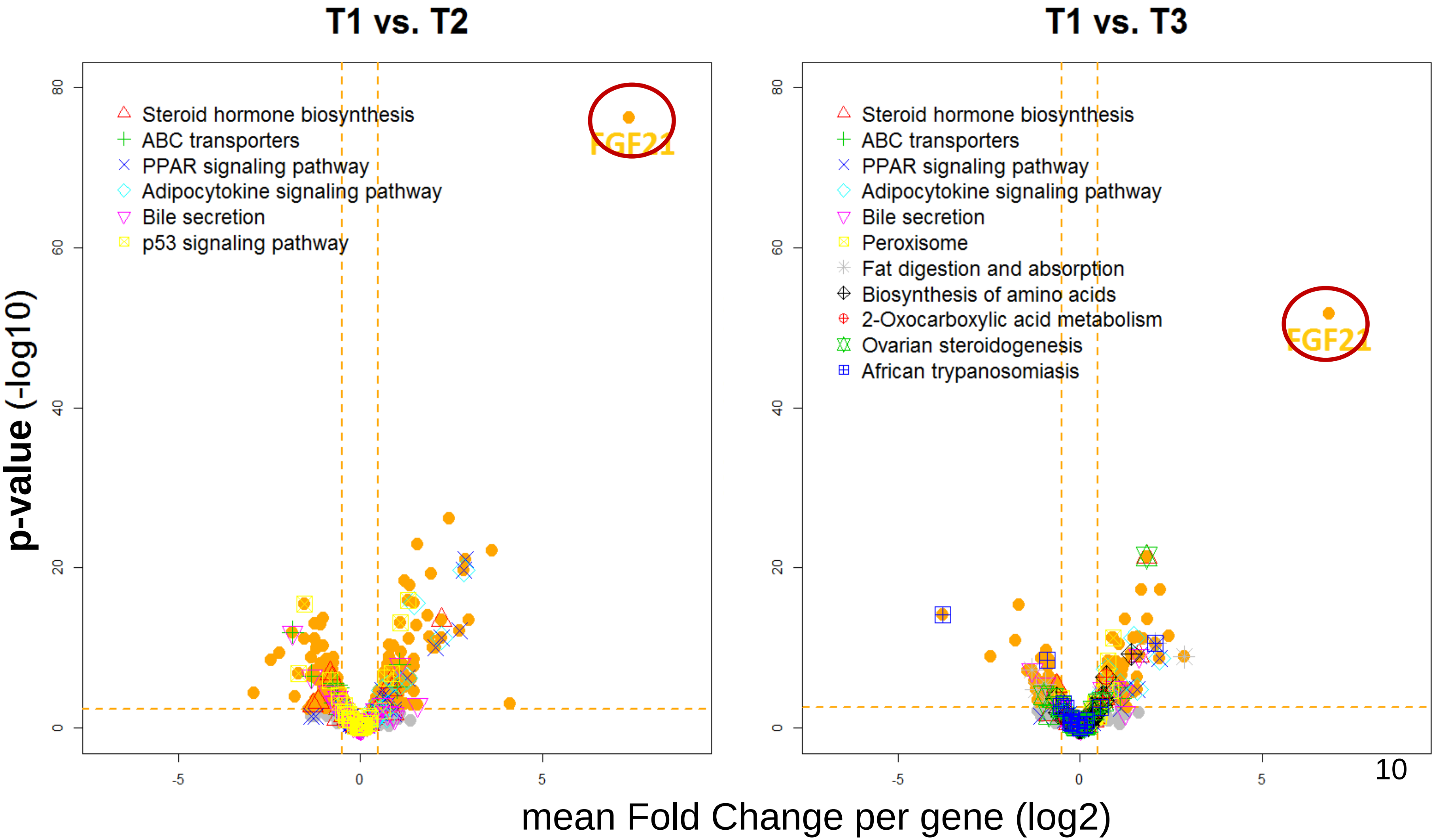
not important

gene in pathway

gene <u>not</u> in pathway

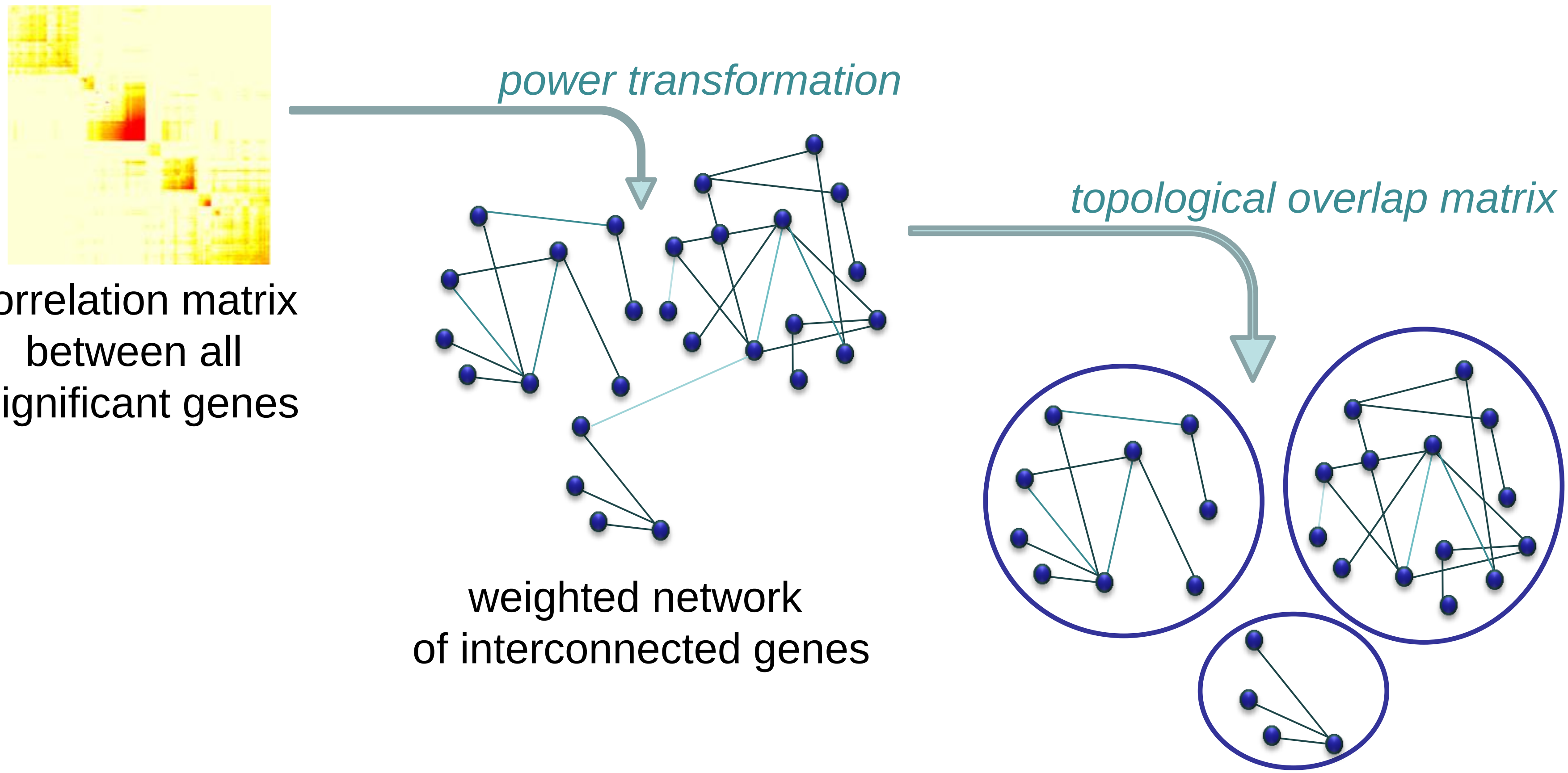
Results – *DGEA and GSEA*

- *~1,000 significant genes (FDR < 5%)*



Methods – *WGCNA*

Weighted Gene Co-expression Network Analysis (WGCNA):



correlation matrix
between all
significant genes

power transformation

weighted network
of interconnected genes

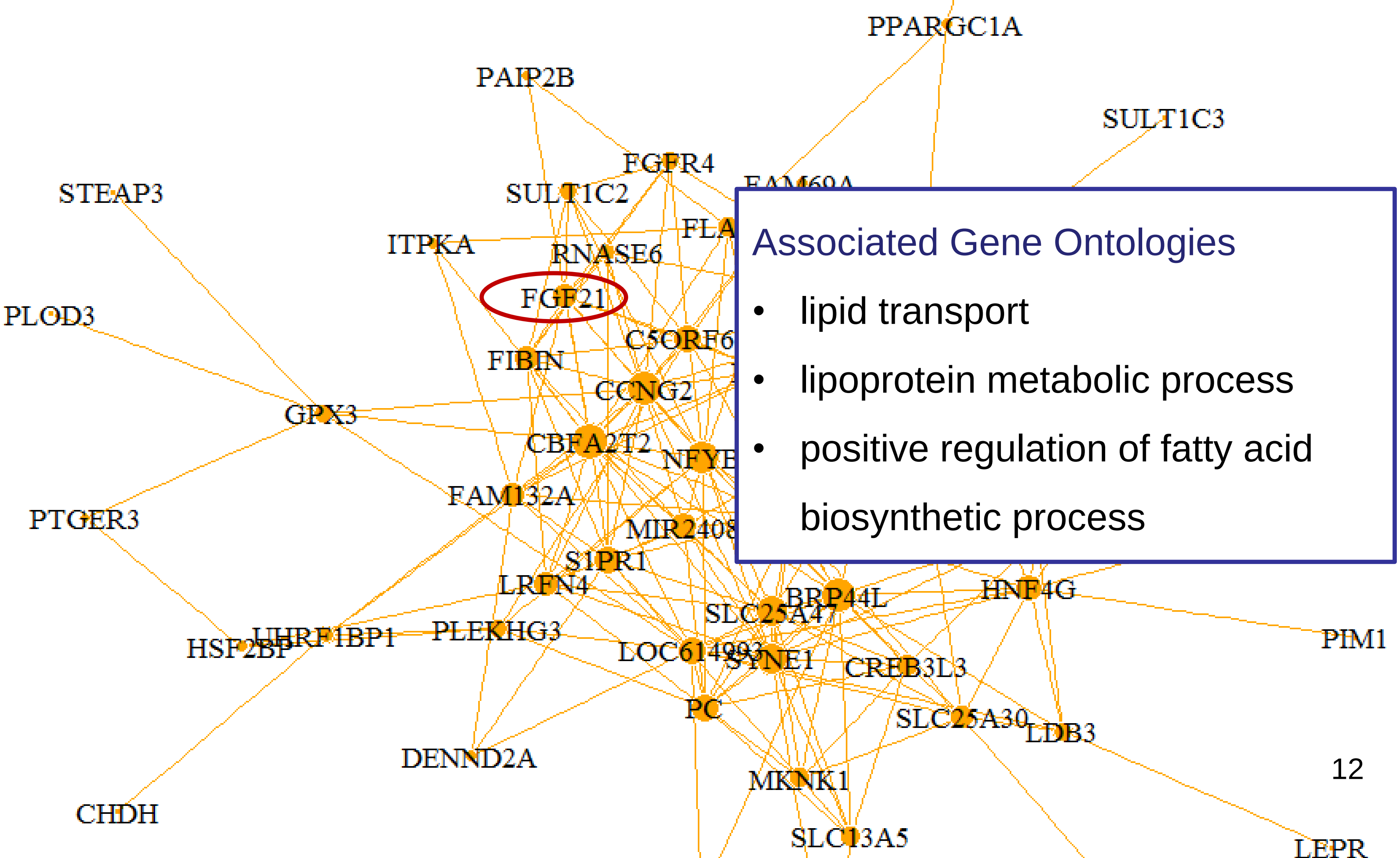
topological overlap matrix

gene modules =
set of tightly connected genes

(Langfelder and Horvarth, 2008)

Results – *WGCNA*

- *15 modules detected (~1,000 significant genes)*



Associated Gene Ontologies

- lipid transport
- lipoprotein metabolic process
- positive regulation of fatty acid biosynthetic process

Conclusions

- ~10% of the ~10,000 genes expressed in the liver are significant comparing ante- vs. post-partum

  ⇨ *importance of liver metabolism for adaptation*

- major hepatic changes involved in gluconeogenesis and lipid mobilization (*PC, FGF21, „adipocytokines signaling pathway“)*

- significant pathways *(e.g. „steroid hormone biosynthesis“)* indicate immunological changes

- identified 15 modules with different expression patterns

  ⇨ *may have different roles in the metabolic adaptation*

Acknowledgements

Josef Johann Gross[1], Cord Drögemüller[2], Fritz Schmitz-Hsu[3], Rupert Bruckmaier[1], Henner Simianer[4]

[1] Veterinary Physiology, Vetsuisse Faculty University of Bern, Switzerland
[2] Institute for Genetics, Vetsuisse Faculty University of Bern, Switzerland
[3] Swissgenetics, Zollikofen, Switzerland
[4] Animal Breeding and Genetics Group, Department of Animal Sciences, Georg-August-University Goettingen, Germany

We especially thank the EAAP for granting a scholarship for the EAAP Meeting 2016 in Belfast.

swissgenetics

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Volkswirtschaftsdepartement EVD
Bundesamt für Berufsbildung und Technologie BBT
Förderagentur für Innovation KTI

This study was supported by a grant of the Swiss Commission for Technology and Innovation CTI (project no. 13948.2 PFLS-LS) and Swissgenetics, Zollikofen, Switzerland.

Thank you for your attention.

References

Drackley, J.K., H.M. Dann, G.N. Douglas, N.A.J. Guretzky, N.B. Litherland, J.P. Underwood, and J.J. Loor. 2005. Physiological and pathological adaptations in dairy cows that may increase susceptibility to periparturient diseases and disorders. *Growth*. 7. doi:10.4081/ijas.2005.323.

Graber, M., S. Kohler, T. Kaufmann, M.G. Doherr, R.M. Bruckmaier, and H.A. van Dorland. 2010. A field study on characteristics and diversity of gene expression in the liver of dairy cows during the transition period. *J. Dairy Sci.* 93:5200–5215. doi:10.3168/jds.2010-3265.

Ingvartsen, K.L., R.J. Dewhurst, and N.C. Friggens. 2003. On the relationship between lactational performance and health: is it yield or metabolic imbalance that cause production diseases in dairy cattle? A position paper. *Livest. Prod. Sci.* 83:277–308. doi:10.1016/S0301-6226(03)00110-6.

Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. 2016. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44:D457-462. doi:10.1093/nar/gkv1070.

Langfelder, P., and S. Horvath. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 9:559. doi:10.1186/1471-2105-9-559.

Robinson, M.D., D.J. McCarthy, and G.K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* 26:139–140. doi:10.1093/bioinformatics/btp616.

Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102:15545–15550. doi:10.1073/pnas.0506580102.

Conclusions

- ~10% of the ~10,000 genes expressed in the liver are significant comparing ante- vs. post-partum
  ⇒ *importance of liver metabolism for adaptation*
- major hepatic changes involved in gluconeogenesis and lipid mobilization (*PC, FGF21, „adipocytokines signaling pathway“)*
- significant pathways *(e.g. „steroid hormone biosynthesis“)* indicate immunological changes
- identified 15 modules with different expression patterns
  ⇒ *may have different roles in the metabolic adaptation*

RNA-Seq

Quality Control and Trimming

FastQC and Trimmomatic

Mapping on reference genome (BT, UMD3.1 rel.85)

STAR:
~ 94% Mapping-Rate
~ 22M Fragmente

Counting reads overlapping exons

featureCounts:
~ 17M Counts

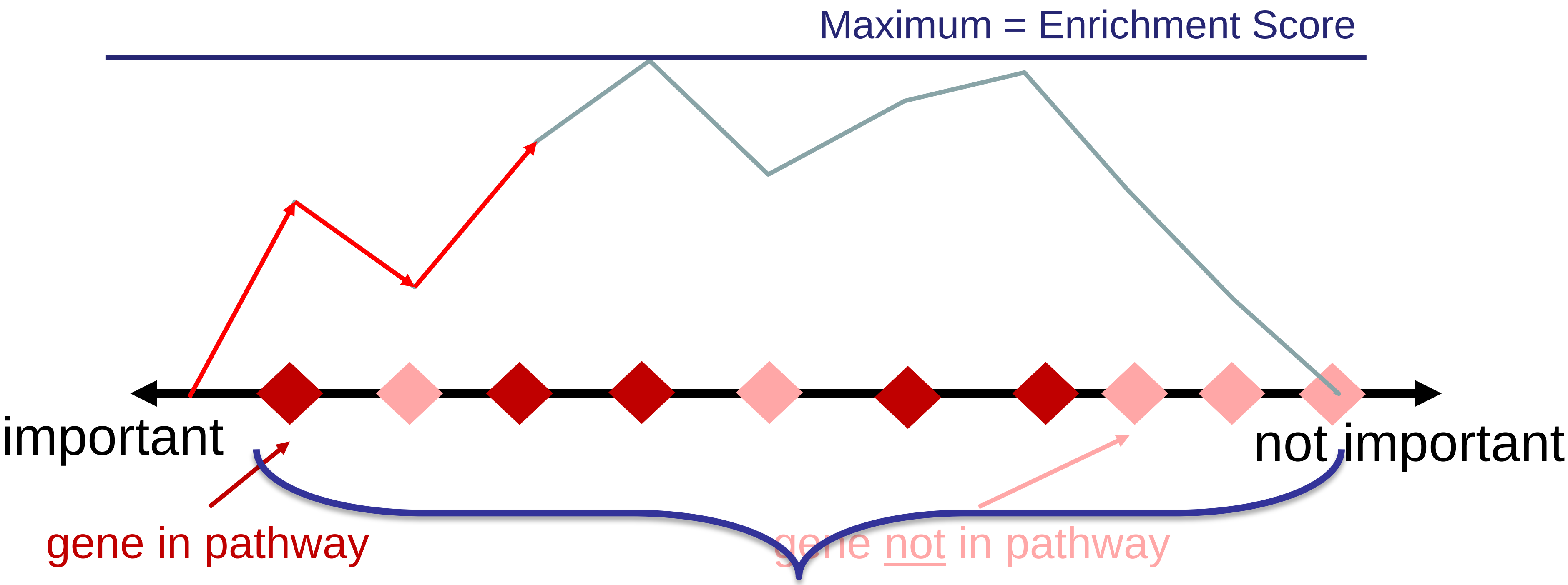
Normalization and modelling of count data

R-Paket edgeR:
TMM, GLM

DGEA, GSEA and WGCNA

GSEA

Gene-Set Enrichment Analysis (Subramanian et al., 2005)



Maximum = Enrichment Score

important

not important

gene in pathway

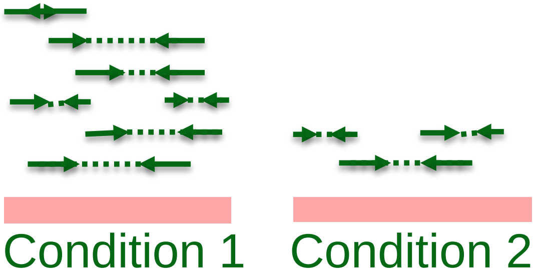
gene not in pathway

$-\log_{10}(p-value)$

Analysis of RNA-Seq Data

Results:

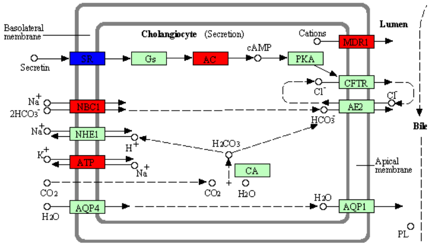
Differential Gene Expression Analysis (DGEA)



⇨

List of differentially expressed genes

Gene Set Enrichment Analysis (GSEA)

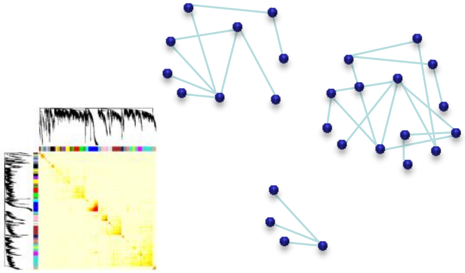


KEGG pathways

⇨

List of differentially expressed pathways

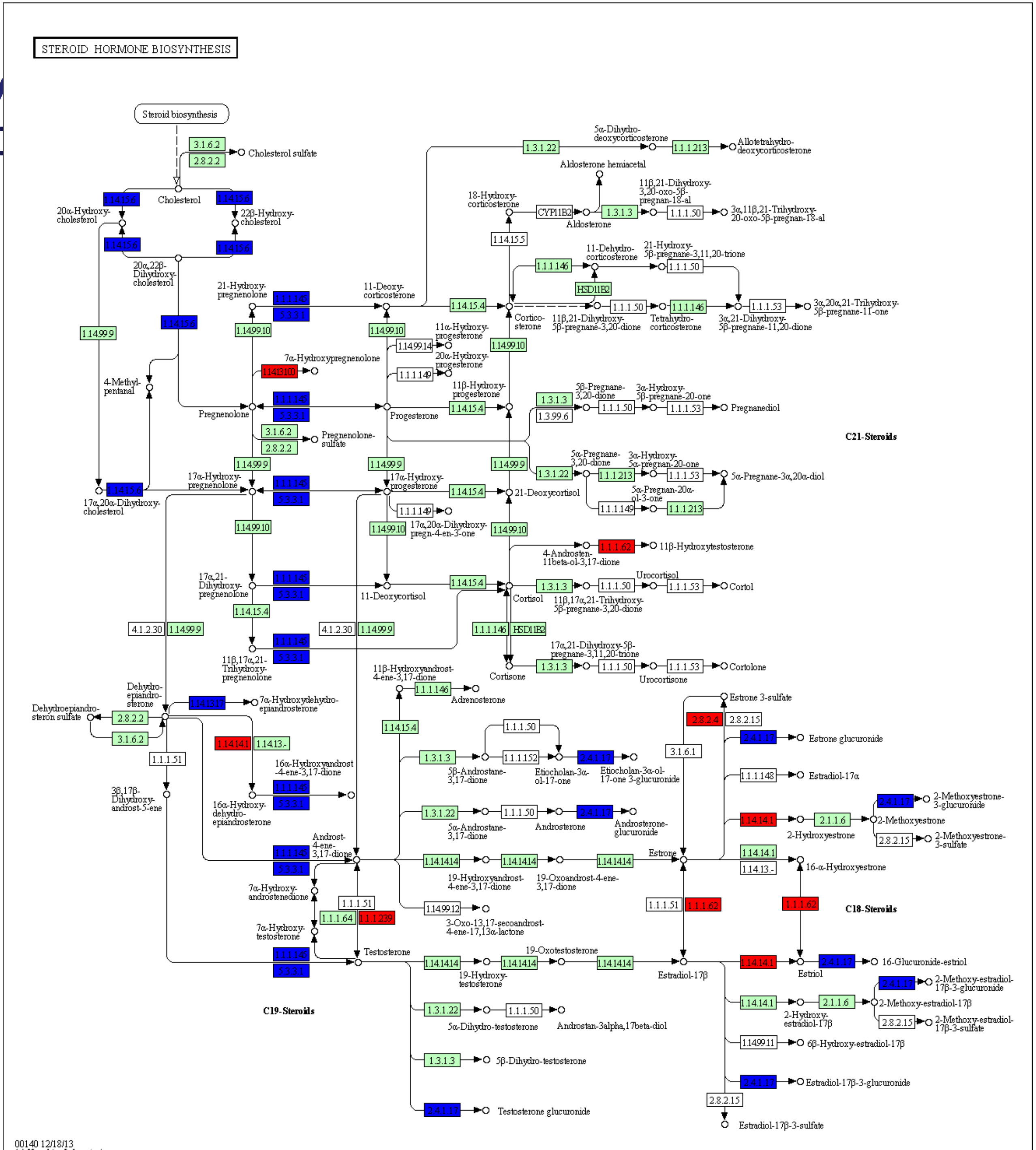
Weighted Gene Co-expression Network Analysis (WGCNA)



⇨

Sets of co-expressed genes differentially expressed

Pathw


STEROID HORMONE BIOSYNTHESIS
Steroid biosynthesis
3.1.6.2
2.8.2.2
Cholesterol sulfate
Cholesterol
1.14.15.6
20α-Hydroxy-cholesterol
22β-Hydroxy-cholesterol
20α,22β-Dihydroxy-cholesterol
1.14.99.9
4-Methyl-pentanal
21-Hydroxy-pregnenolone
1.1.1.145
5.3.3.1
1.14.99.10
7α-Hydroxypregnenolone
1.14.13.100
Pregnenolone
Pregnenolone-sulfate
17α-Hydroxy-pregnenolone
17α,20α-Dihydroxy-cholesterol
17α,21-Dihydroxy-pregnenolone
1.14.15.4
11β,17α,21-Trihydroxy-pregnenolone
4.1.2.30
11-Deoxy-corticosterone
11α-Hydroxy-progesterone
1.14.99.14
20α-Hydroxy-progesterone
1.1.1.149
Progesterone
17α-Hydroxy-progesterone
17α,20α-Dihydroxy-pregn-4-en-3-one
11-Deoxycortisol
11β-Hydroxy-progesterone
21-Deoxycortisol
5α-Dihydro-deoxycorticosterone
1.3.1.22
1.1.1.213
Allotetrahydro-deoxycorticosterone
Aldosterone hemiacetal
18-Hydroxy-corticosterone
CYP11B2
1.3.1.3
Aldosterone
11β,21-Dihydroxy-3,20-oxo-5β-pregnan-18-al
1.1.1.50
3α,11β,21-Trihydroxy-20-oxo-5β-pregnan-18-al
1.14.15.5
11-Dehydro-corticosterone
1.1.1.146
21-Hydroxy-5β-pregnane-3,11,20-trione
HSD11B2
Cortico-sterone
11β,21-Dihydroxy-5β-pregnane-3,20-dione
Tetrahydro-corticosterone
3α,21-Dihydroxy-5β-pregnane-11,20-dione
1.1.1.53
3α,20α,21-Trihydroxy-5β-pregnane-11-one
5β-Pregnane-3,20-dione
1.3.99.6
3α-Hydroxy-5β-pregnane-20-one
Pregnanediol
C21-Steroids
5α-Pregnane-3,20-dione
3α-Hydroxy-5α-pregnan-20-one
5α-Pregnane-3α,20α-diol
5α-Pregnan-20α-ol-3-one
4-Androsten-11beta-ol-3,17-dione
1.1.1.62
11β-Hydroxytestosterone
Cortisol
Urocortisol
Cortol
11β,17α,21-Trihydroxy-5β-pregnane-3,20-dione
17α,21-Dihydroxy-5β-pregnane-3,11,20-trione
Cortisone
Urocortisone
Cortolone
11β-Hydroxyandrost-4-ene-3,17-dione
Adrenosterone
Dehydro-epiandro-sterone
Dehydroepiandro-steron sulfate
1.14.13.17
7α-Hydroxydehydro-epiandrosterone
1.14.14.1
1.14.13.-
1.1.1.51
3β,17β-Dihydroxy-androst-5-ene
16α-Hydroxyandrost-4-ene-3,17-dione
16α-Hydroxy-dehydro-epiandrosterone
Androst-4-ene-3,17-dione
7α-Hydroxy-androstenedione
7α-Hydroxy-testosterone
1.1.1.64
1.1.1.239
Testosterone
C19-Steroids
5β-Androstane-3,17-dione
1.1.1.152
Etiocholan-3α-ol-17-one
2.4.1.17
Etiocholan-3α-ol-17-one 3-glucuronide
5α-Androstane-3,17-dione
Androsterone
Androsterone-glucuronide
1.14.14.14
19-Hydroxyandrost-4-ene-3,17-dione
19-Oxoandrost-4-ene-3,17-dione
1.14.99.12
3-Oxo-13,17-secoandrost-4-ene-17,13α-lactone
19-Hydroxy-testosterone
19-Oxotestosterone
Estradiol-17β
5α-Dihydro-testosterone
Androstan-3alpha,17beta-diol
5β-Dihydro-testosterone
Testosterone glucuronide
Estrone 3-sulfate
2.8.2.4
2.8.2.15
3.1.6.1
Estrone glucuronide
1.1.1.148
Estradiol-17α
2.1.1.6
2-Hydroxyestrone
2-Methoxyestrone
2-Methoxyestrone-3-glucuronide
2-Methoxyestrone-3-sulfate
Estrone
16-α-Hydroxyestrone
C18-Steroids
Estriol
16-Glucuronide-estriol
2-Hydroxy-estradiol-17β
2-Methoxy-estradiol-17β
2-Methoxy-estradiol-17β-3-glucuronide
2-Methoxy-estradiol-17β-3-sulfate
1.14.99.11
6β-Hydroxy-estradiol-17β
Estradiol-17β-3-glucuronide
Estradiol-17β-3-sulfate
00140 12/18/13

Mappen der Reads – STAR vs Tophat2

Reference genome und gene annotation:

- *Ensembl UMD3.1 Release 85*
- *24‘616 genes (19‘994 coding)*

STAR	**Tophat2**
• fast (~20 min)	• slow (~12 h)
• suggested by *GATK Best Practices*	• high mapping rate (Kim et. al., 2013)
• high mapping rate (Engström et al., 2013)	• high rate of false positives (Dobin and Gingeras, 2013)
• *Mapping-Rate: ~94%*	• *Mapping-Rate: ~81%*

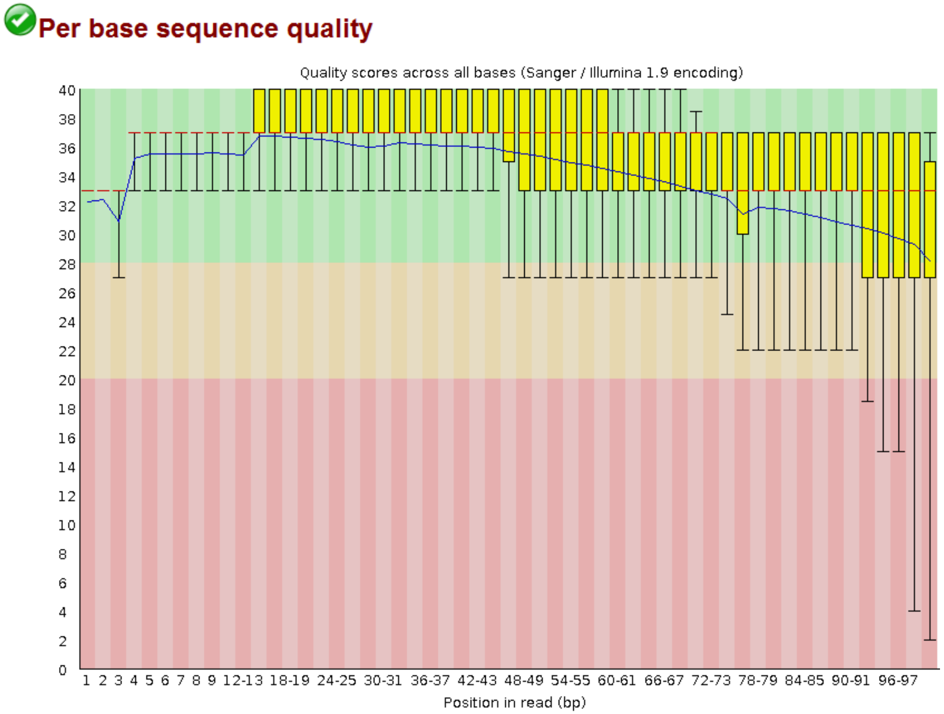
In our study minor differences

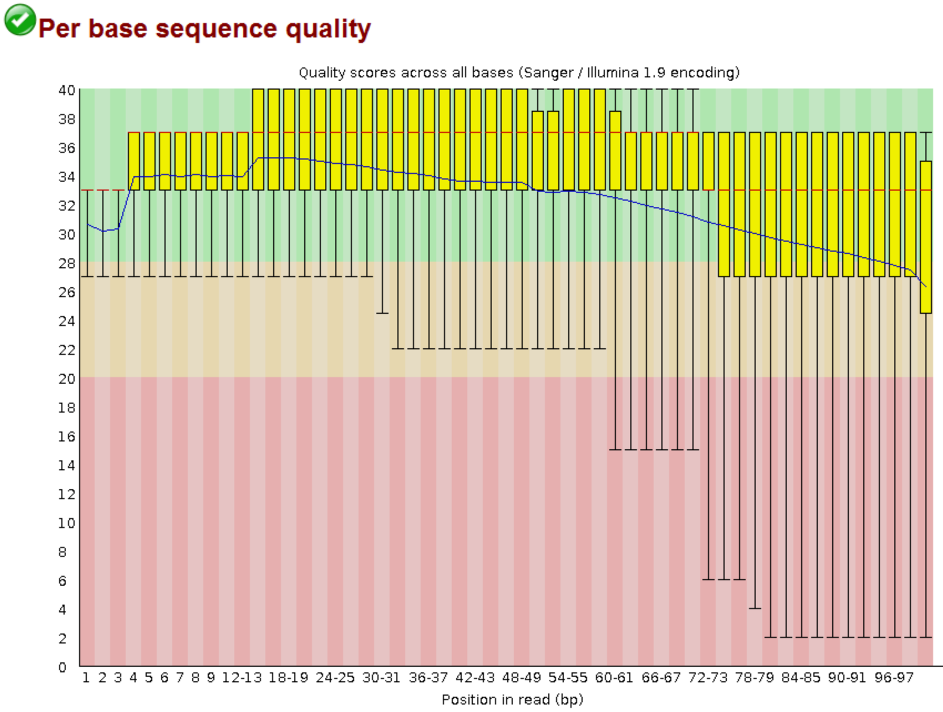
Qualitätskontrolle – FastQC

RNA fragment

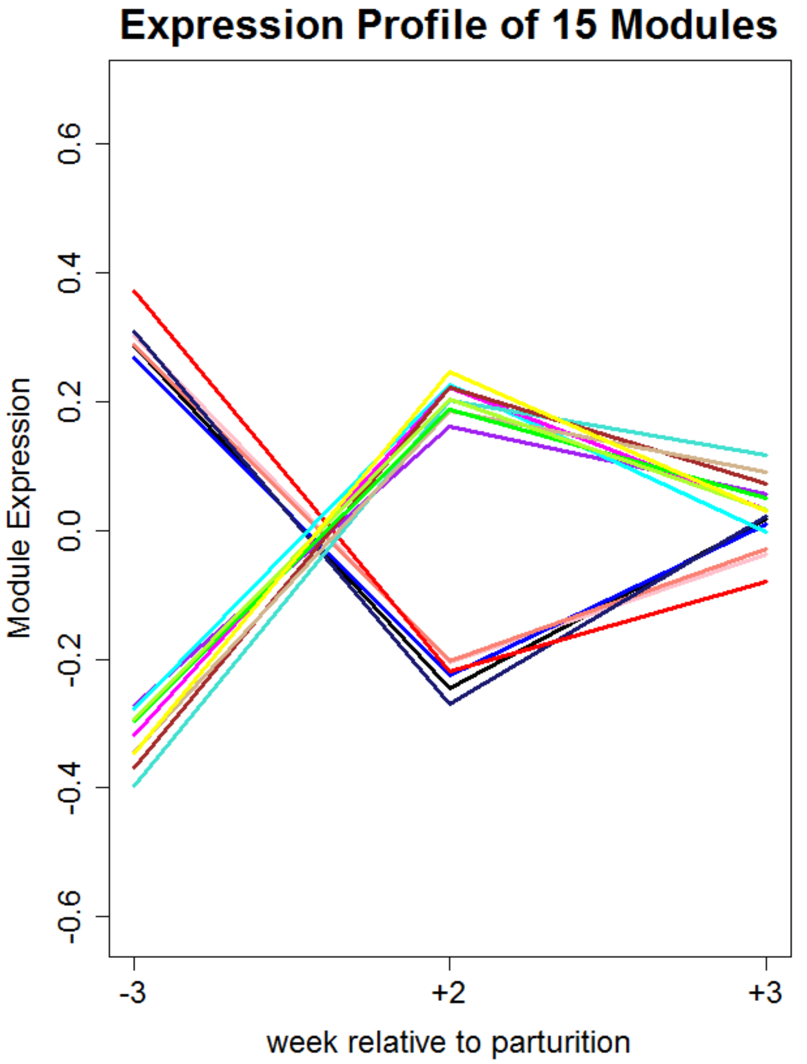
Forward Read: 101 bp

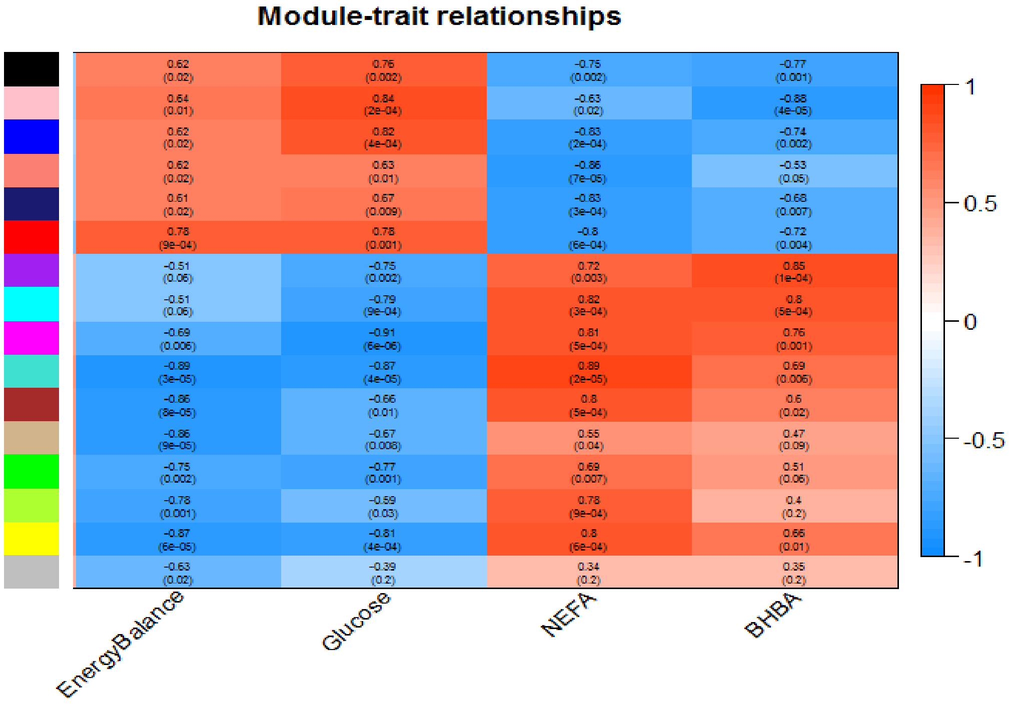
Reverse Read: 101 bp





Results





	EnergyBalance	Glucose	NEFA	BHBA
⬛	0.62 (0.02)	0.76 (0.002)	−0.75 (0.002)	−0.77 (0.001)
🟥	0.64 (0.01)	0.84 (2e−04)	−0.63 (0.02)	−0.88 (4e−05)
🟦	0.62 (0.02)	0.82 (4e−04)	−0.83 (2e−04)	−0.74 (0.002)
🟥	0.62 (0.02)	0.63 (0.01)	−0.86 (7e−05)	−0.53 (0.05)
🟦	0.61 (0.02)	0.67 (0.009)	−0.83 (3e−04)	−0.68 (0.007)
🟥	0.78 (9e−04)	0.78 (0.001)	−0.8 (6e−04)	−0.72 (0.004)
🟪	−0.51 (0.06)	−0.75 (0.002)	0.72 (0.003)	0.85 (1e−04)
🟦	−0.51 (0.06)	−0.79 (9e−04)	0.82 (3e−04)	0.8 (5e−04)
🟪	−0.69 (0.006)	−0.91 (6e−06)	0.81 (5e−04)	0.76 (0.001)
🟩	−0.89 (3e−05)	−0.87 (4e−05)	0.89 (2e−05)	0.69 (0.006)
🟫	−0.86 (8e−05)	−0.66 (0.01)	0.8 (5e−04)	0.6 (0.02)
🟫	−0.86 (9e−05)	−0.67 (0.008)	0.55 (0.04)	0.47 (0.09)
🟩	−0.75 (0.002)	−0.77 (0.001)	0.69 (0.007)	0.51 (0.06)
🟩	−0.78 (0.001)	−0.59 (0.03)	0.78 (9e−04)	0.4 (0.2)
🟨	−0.87 (6e−05)	−0.81 (4e−04)	0.8 (6e−04)	0.66 (0.01)
⬜	−0.63 (0.02)	−0.39 (0.2)	0.34 (0.2)	0.35 (0.2)