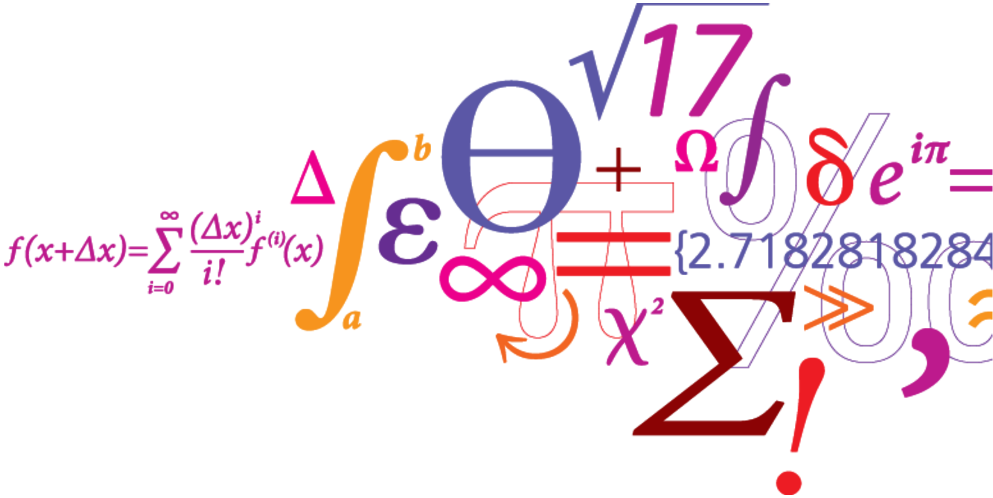
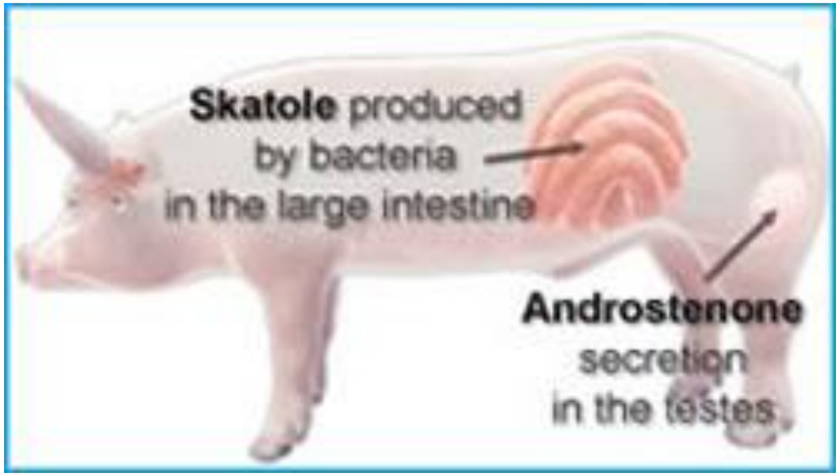
Genome-wide DNA methylation analysis reveals candidate epigenetic biomarkers of boar taint in pigs

Xiao Wang[1,2], Markus Drag[3] and Haja N. Kadarmideen[1,2]
[1]Technical University of Denmark, DTU Bioinformatics, Lyngby, Denmark
[2]Technical University of Denmark, DTU Compute, Lyngby, Denmark
[3]University of Copenhagen, Department of Veterinary and Animal Sciences, Frederiksberg, Denmark



Background

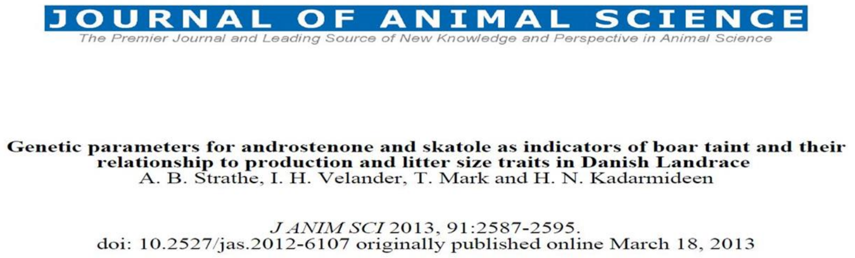
- The offensive odor of boar taint (BT) is primarily caused by the accumulation of skatole and androstenone
- Skatole and androstenone traits with high heritability (0.33 and 0.59)
- Selection of low BT boars can be an effective approach to avoid BT and other disadvantages of surgical castration



Patterson, R. L. S. (1968), Gower, D. B. (1972), Strathe, A. B. *et al*., (2013)

Genetics and transcriptomics of boar taint

- **Earlier published work in boar taint from our group - genetic parameter estimation, genomic selection and transcriptomics**



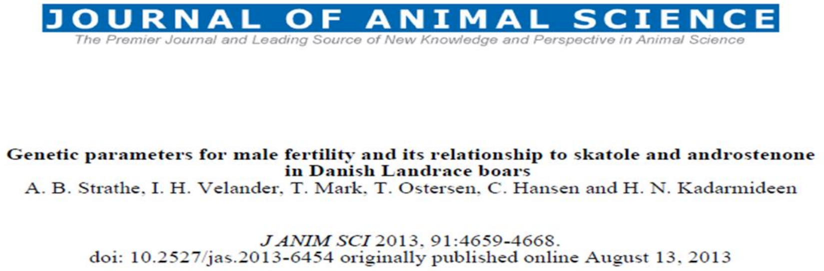
JOURNAL OF ANIMAL SCIENCE

The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science

Genetic parameters for androstenone and skatole as indicators of boar taint and their relationship to production and litter size traits in Danish Landrace

A. B. Strathe, I. H. Velander, T. Mark and H. N. Kadarmideen

J ANIM SCI 2013, 91:2587-2595.
doi: 10.2527/jas.2012-6107 originally published online March 18, 2013



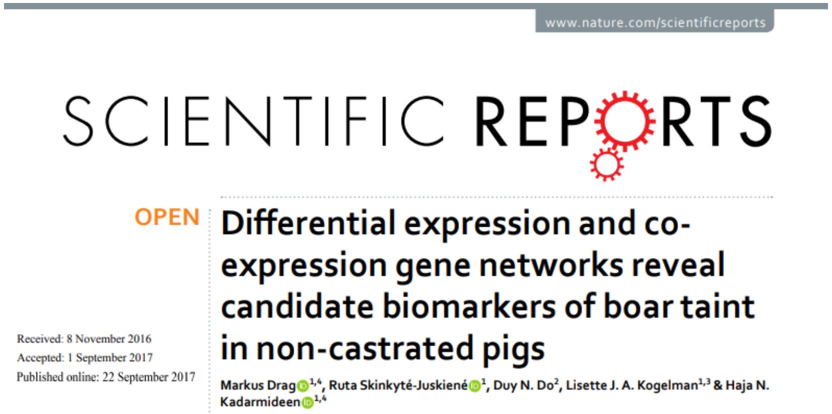
JOURNAL OF ANIMAL SCIENCE

The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science

Genetic parameters for male fertility and its relationship to skatole and androstenone in Danish Landrace boars

A. B. Strathe, I. H. Velander, T. Mark, T. Ostersen, C. Hansen and H. N. Kadarmideen

J ANIM SCI 2013, 91:4659-4668.
doi: 10.2527/jas.2013-6454 originally published online August 13, 2013



www.nature.com/scientificreports

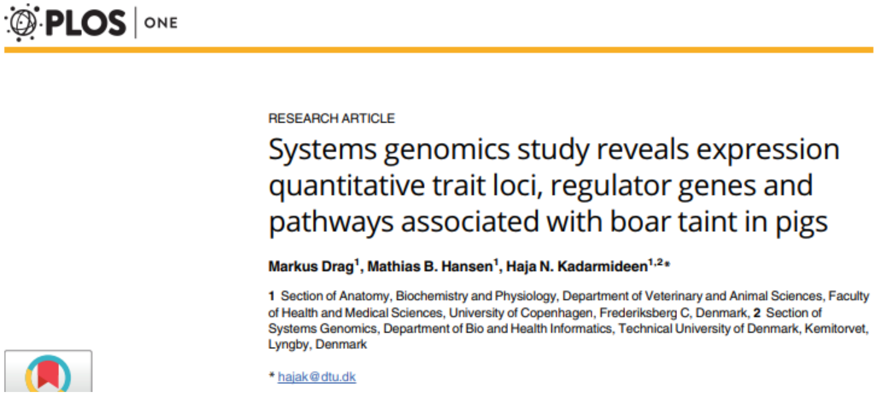
SCIENTIFIC REPORTS

OPEN

Differential expression and co-expression gene networks reveal candidate biomarkers of boar taint in non-castrated pigs

Received: 8 November 2016
Accepted: 1 September 2017
Published online: 22 September 2017

Markus Drag[1,4], Ruta Skinkyté-Juskiené[1], Duy N. Do[2], Lisette J. A. Kogelman[1,3] & Haja N. Kadarmideen[1,4]



PLOS | ONE

RESEARCH ARTICLE

Systems genomics study reveals expression quantitative trait loci, regulator genes and pathways associated with boar taint in pigs

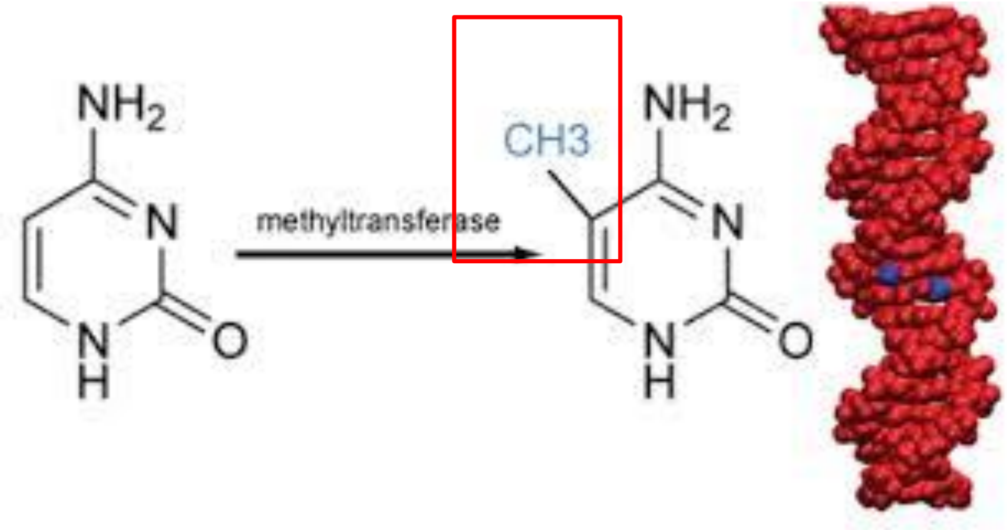
Markus Drag[1], Mathias B. Hansen[1], Haja N. Kadarmideen[1,2]*

1 Section of Anatomy, Biochemistry and Physiology, Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Frederiksberg C, Denmark, 2 Section of Systems Genomics, Department of Bio and Health Informatics, Technical University of Denmark, Kemitorvet, Lyngby, Denmark

* hajak@dtu.dk

Epigenetics of boar taint

- Epigenetics is defined as changes in gene function that are heritable and no change in DNA sequence
- DNA methylation has been examined to be associated with growth, immune response and reproduction traits in pigs

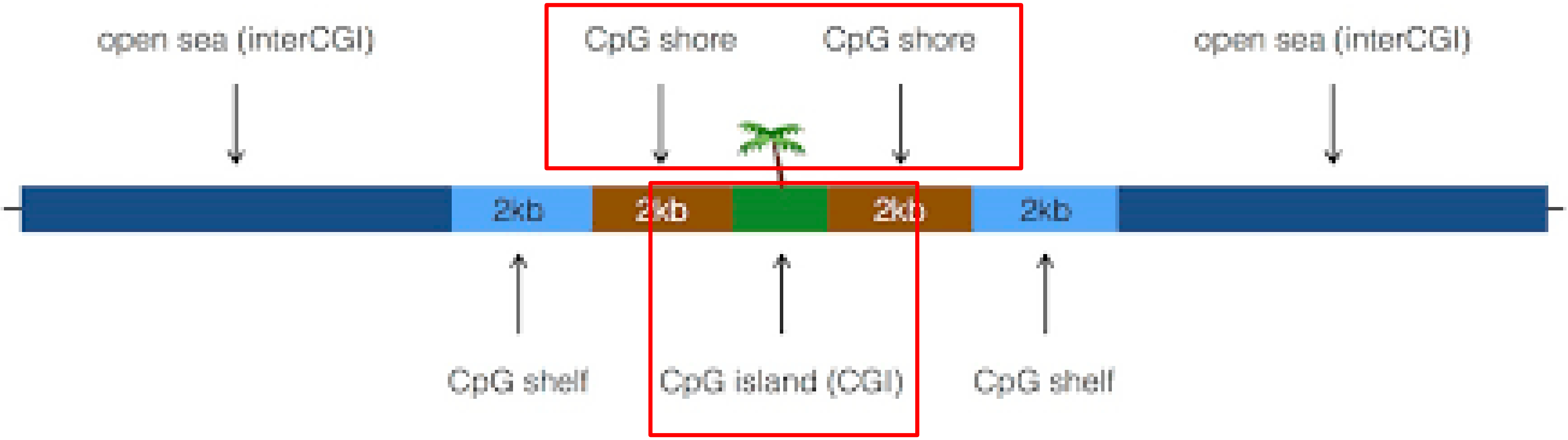


Wu, C. T. and J. R. Morris. (2001)

CpG island

- CpG islands were defined as a region with at least 200 bp, a GC fraction more than 0.5 and an observed-to-expected ratio of CpG more than 0.6

- CpG island shores were defined as regions 2 kb in length adjacent to CpG islands

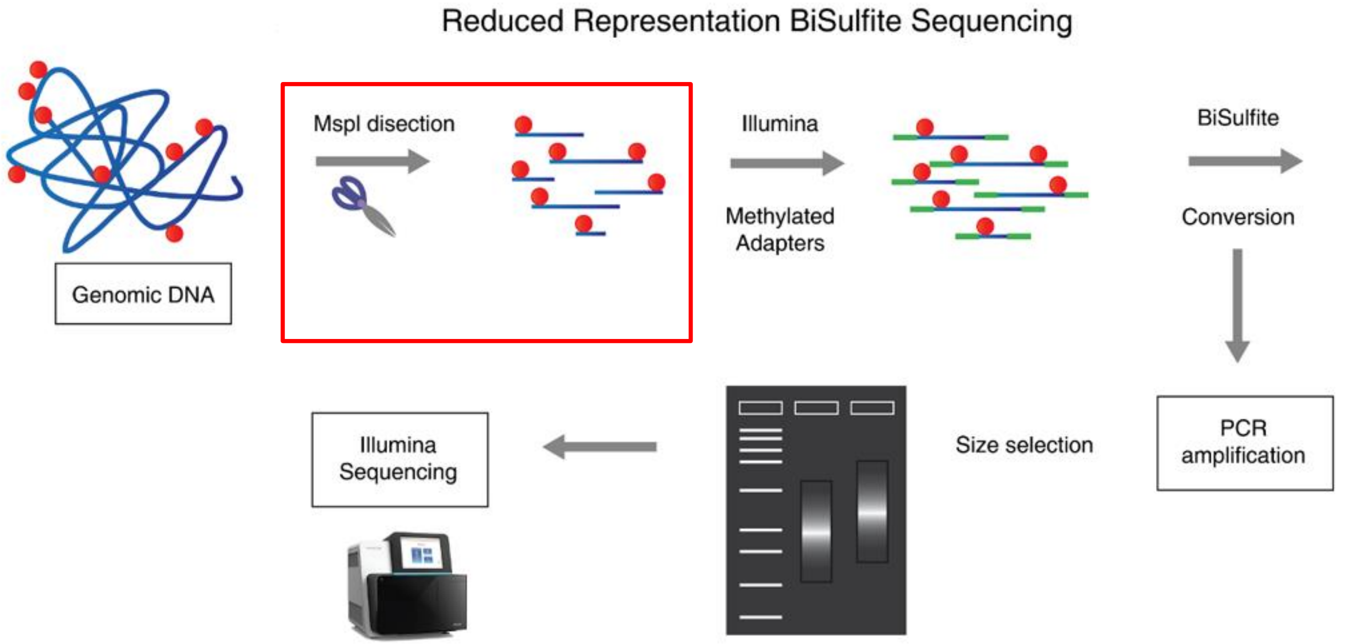


Gardiner-Garden, M. and Frommer, M. (1987)

Materials

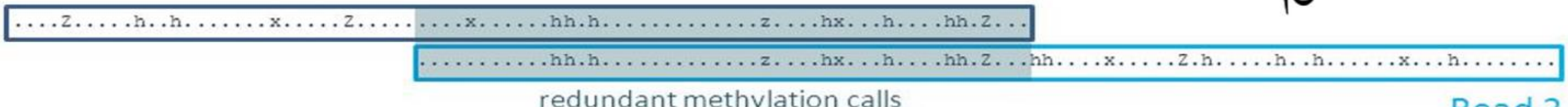
$$Summarised\ BT\ EBV = Skatole\ EBV + Human\ nose\ score\ EBV$$

Low BT EBV	**3 testis sample**	**9 RRBS (Reduced representation bisulfite sequencing)**
Medium BT EBV	**3 testis sample**	
High BT EBV	**3 testis sample**	

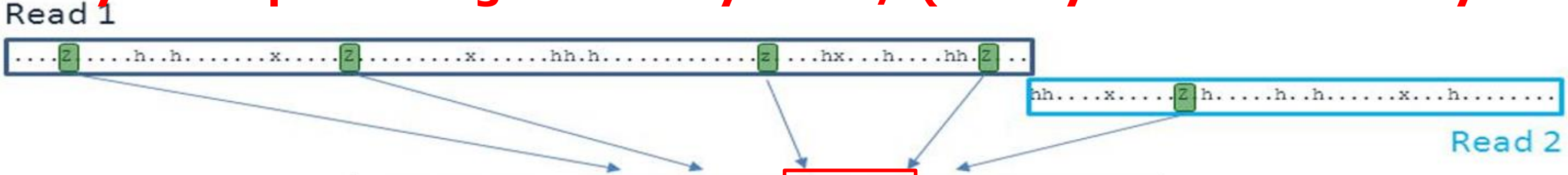


Methylation extraction

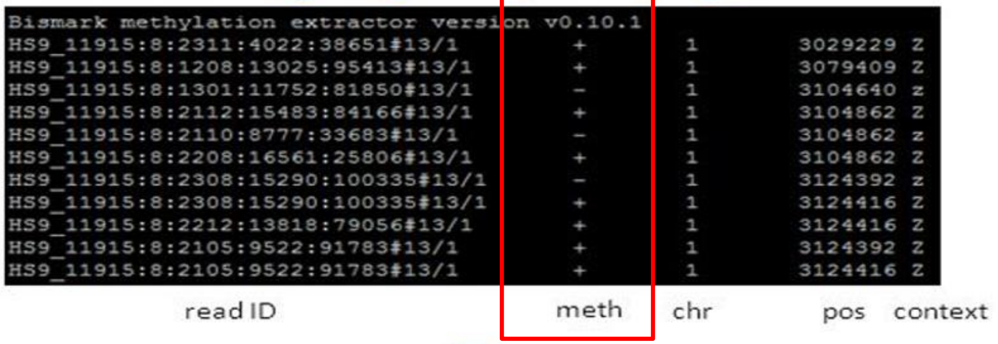
Read 1



Methylation percentage = methylated / (methylated + unmethylated)



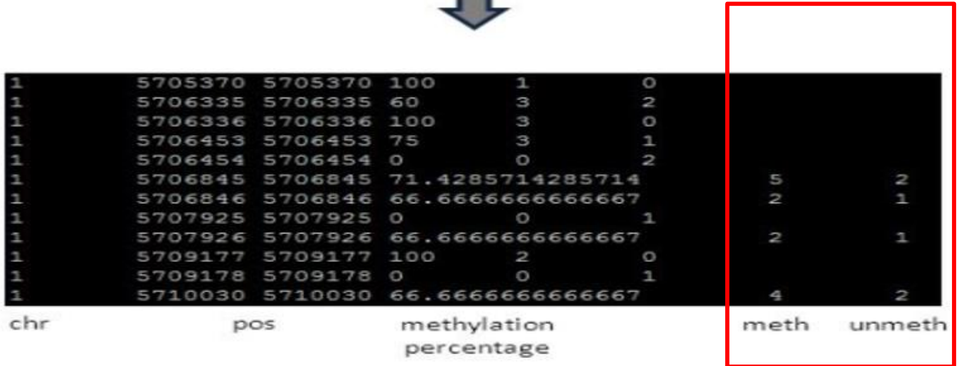
CpG methylation output



```
Bismark methylation extractor version v0.10.1
HS9_11915:8:2311:4022:38651#13/1         +    1        3029229 Z
HS9_11915:8:1208:13025:95413#13/1        +    1        3079409 Z
HS9_11915:8:1301:11752:81850#13/1        -    1        3104640 z
HS9_11915:8:2112:15483:84166#13/1        +    1        3104862 Z
HS9_11915:8:2110:8777:33683#13/1         -    1        3104862 z
HS9_11915:8:2208:16561:25806#13/1        +    1        3104862 Z
HS9_11915:8:2308:15290:100335#13/1       -    1        3124392 z
HS9_11915:8:2308:15290:100335#13/1       +    1        3124416 Z
HS9_11915:8:2212:13818:79056#13/1        +    1        3124416 Z
HS9_11915:8:2105:9522:91783#13/1         +    1        3124392 Z
HS9_11915:8:2105:9522:91783#13/1         +    1        3124416 Z
```

read ID meth chr pos context

coverage output



```
1       5705370 5705370 100     1       0
1       5706335 5706335 60      3       2
1       5706336 5706336 100     3       0
1       5706453 5706453 75      3       1
1       5706454 5706454 0       0       2
1       5706845 5706845 71.4285714285714        5       2
1       5706846 5706846 66.6666666666667        2       1
1       5707925 5707925 0       0       1
1       5707926 5707926 66.6666666666667        2       1
1       5709177 5709177 100     2       0
1       5709178 5709178 0       0       1
1       5710030 5710030 66.6666666666667        4       2
```

chr pos methylation percentage meth unmeth

Babraham Bioinformatics

Methods - DMC

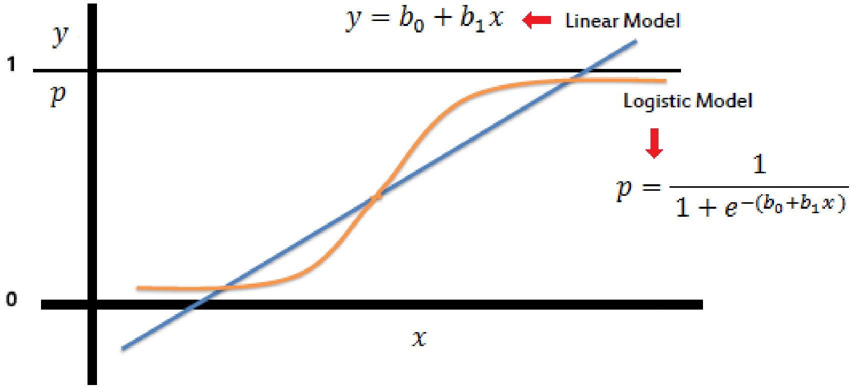
- Differentially methylated cytosine (DMC) using methylKit package through the logistic regression model:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 T_i,$$

where π_i is the methylation proportion at a cytosine , and T_i is the treatment indicator (high or medium or low BT level)

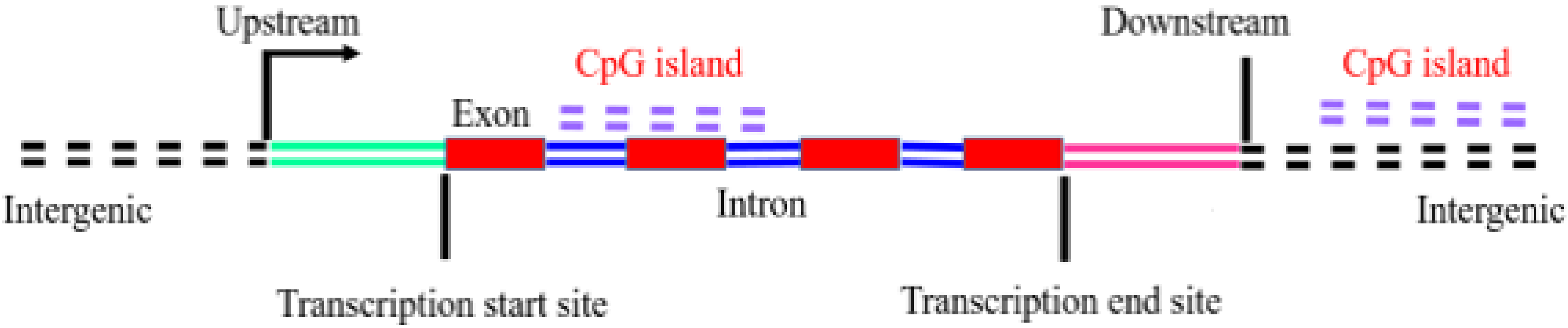
- *P*-values were calculated and then adjusted to Q-values using false discovery rate (FDR) to correct multiple testing

Akalin, A. *et al.*, (2012).



Methods - annotation

- **DMCs were annotated within a 10 kb upstream region from the nearest transcription start site (TSS), exonic, intronic and intergenic regions**



- **Differentially expressed (DE) analysis in Drag's study**
- **DE genes from Gene Expression Omnibus (GEO) by FDR < 0.01**

Drag, M. *et al*., (2017)

Technical flow

Raw RRBS

Adapters Remove + quality control (Trimmomatic software)

Clean data of RRBS

Mapping (Bismark Bisulfite Mapper)

Mapped RRBS statistics

Differentially methylated cytosine (DMC) identification (R package methylKit)

DMC analysis

DMC annotation (R package genomation)

Associated gene analysis

Go + pathway (GenCLiP 2.0 software)

Results - mapping

Uniquely aligned rate: 49%

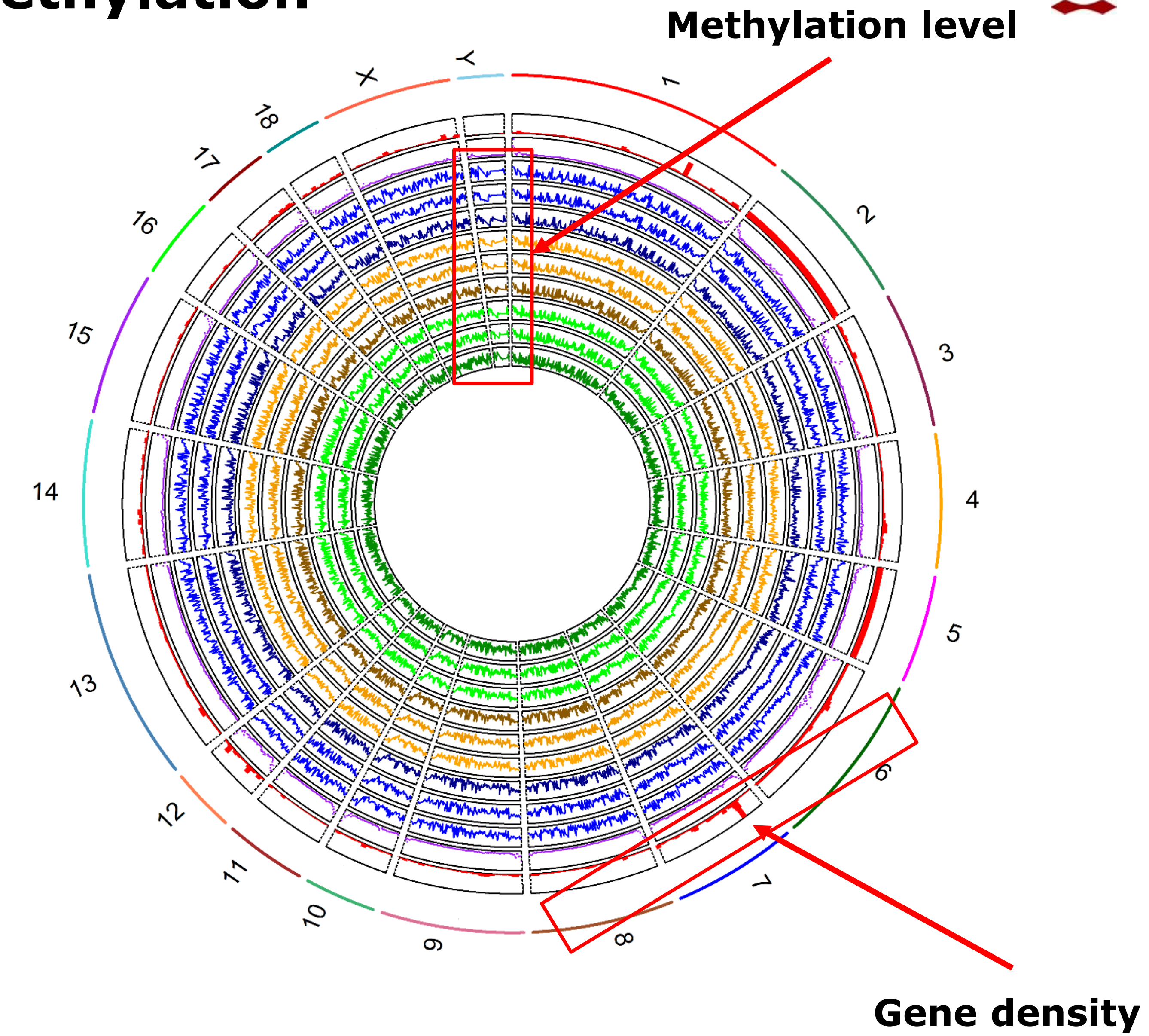
CpG methylation rate: 46% to 53%

Sample ID	Clean read pairs	Uniquely aligned rate	Number of aligned sites	Total number of analyzed cytosine	Cytosine methylation rate in CpG context	Cytosine methylation rate in CHG context	Cytosine methylation rate in CHH context
Low 1	16,505,578	46%	6,555,417	210,492,580	49%	0.91%	0.61%
Low 2	93,817,089	51%	11,786,693	1,458,034,594	53%	0.99%	0.69%
Low 3	38,026,074	47%	8,350,750	507,968,318	46%	0.84%	0.58%
Medium 1	75,769,839	51%	11,024,632	1,161,664,236	52%	0.87%	0.62%
Medium 2	57,267,890	51%	10,230,855	994,282,472	50%	0.68%	0.52%
Medium 3	68,607,455	46%	8,427,406	881,065,710	46%	0.89%	0.64%
High 1	85,068,927	49%	8,799,356	1,220,798,901	49%	0.92%	0.67%
High 2	75,438,276	51%	9,259,657	1,194,394,820	51%	0.92%	0.67%
High 3	16,940,690	47%	6,619,706	214,465,154	50%	0.95%	0.66%

Results - methylation

- **Variation between biological replicates was low**

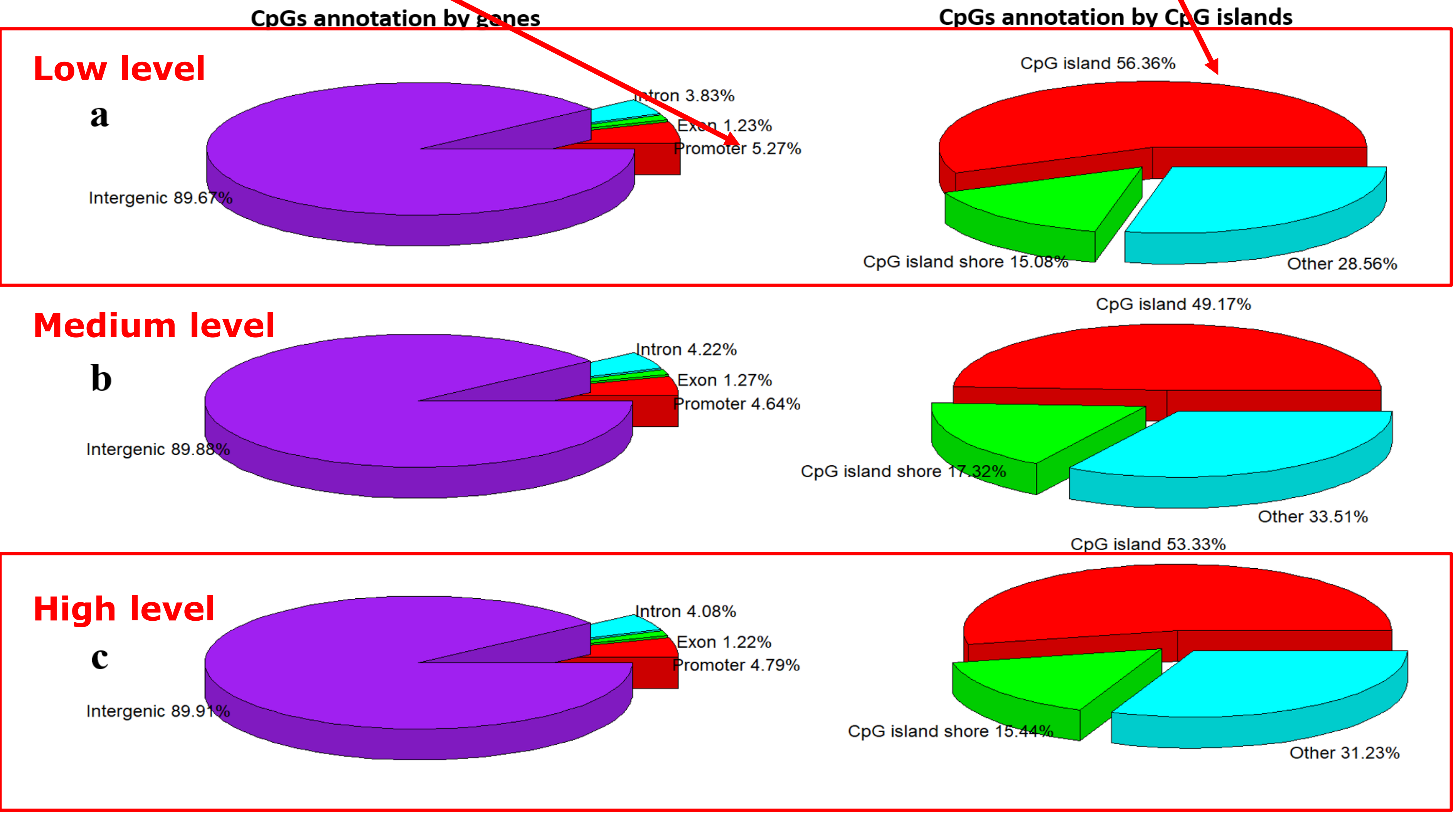
- **Coefficient of density of genes regression on methylation level: -2.2 ($P < 0.001$)**



Results - annotation

Promoter: 4.64% ~ 5.27%

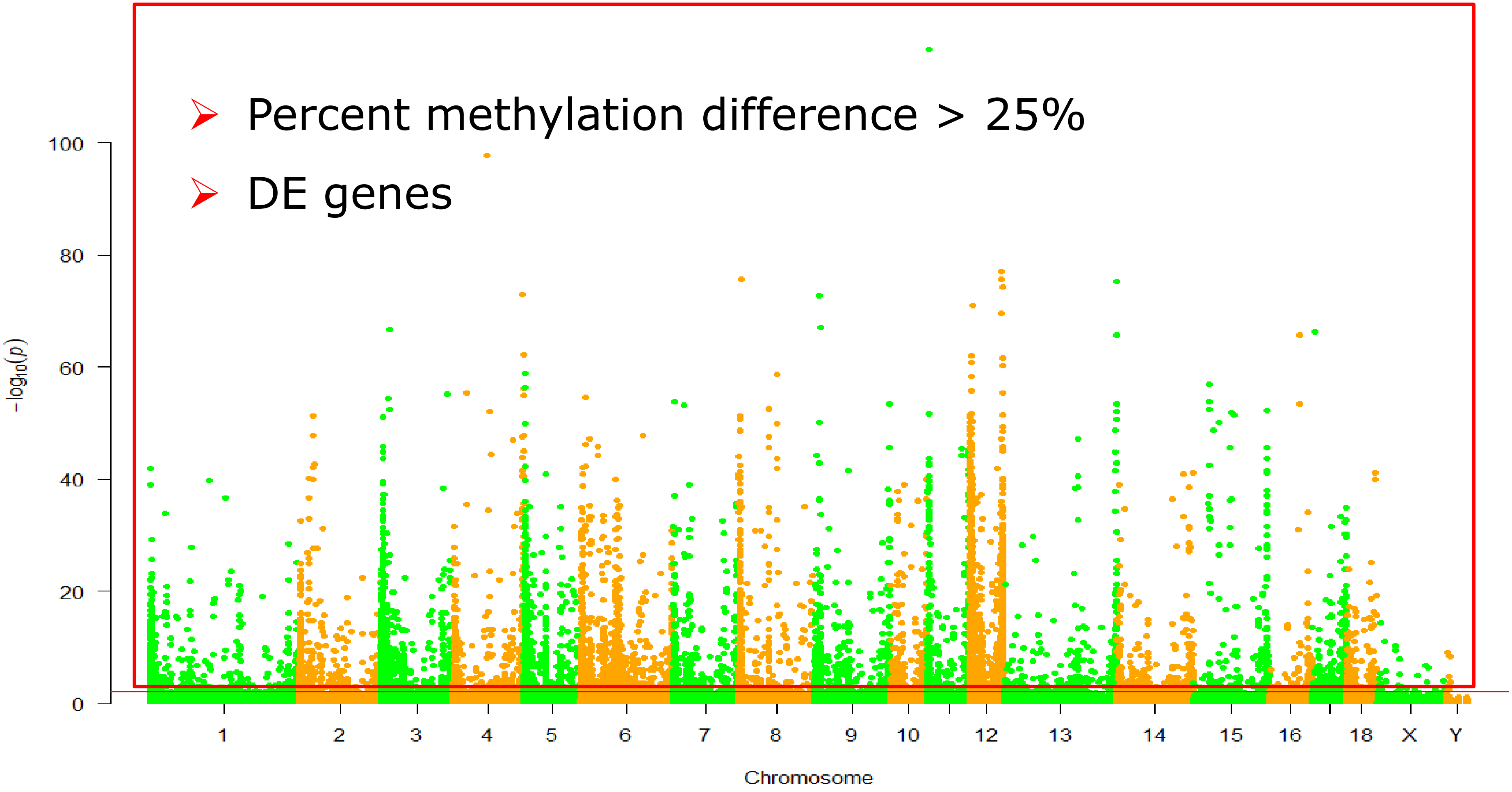
CpG islands: 49.17% ~ 56.36%



Results - DMC

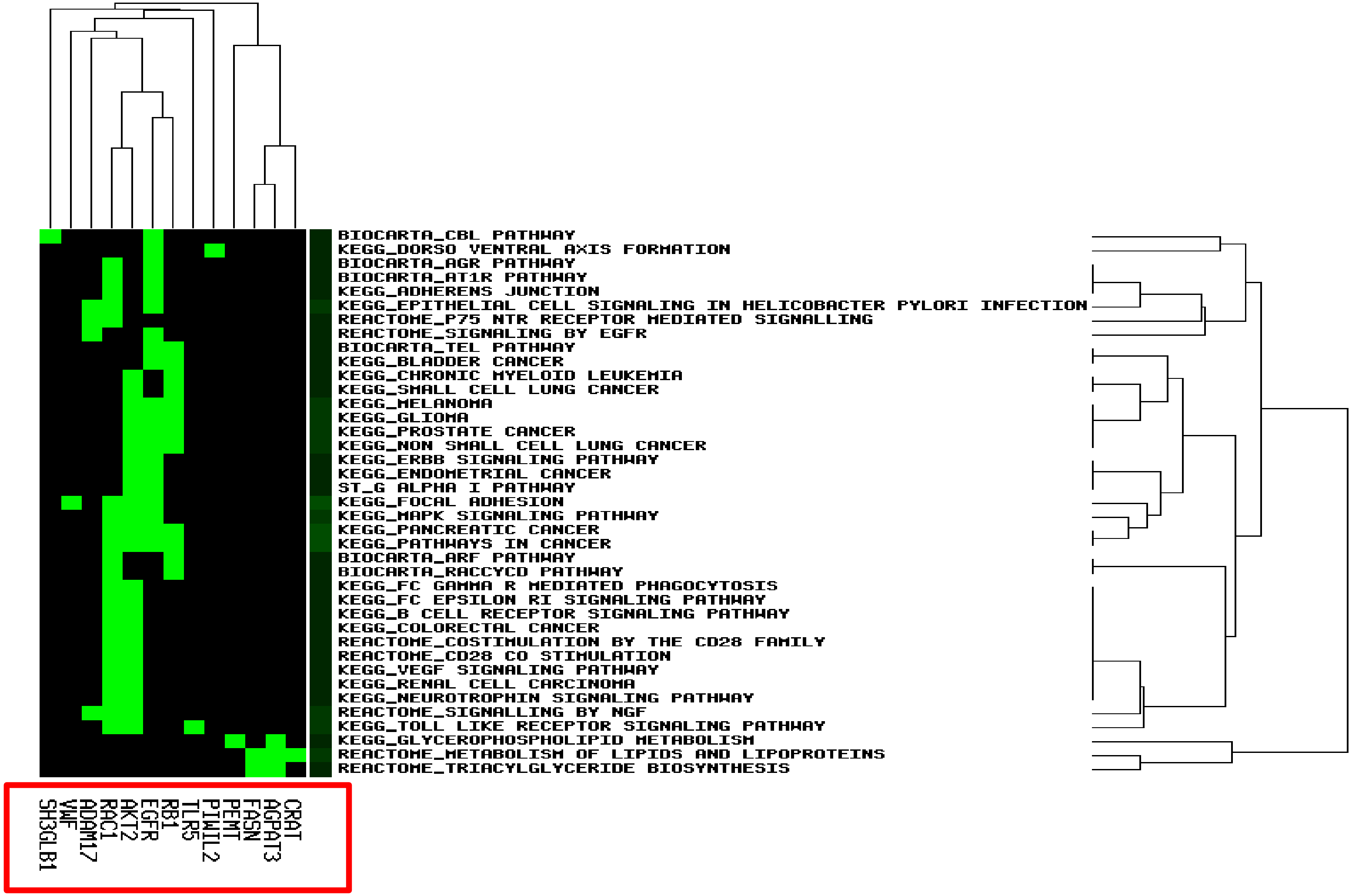
Promoter	Exon	Intron	Intergenic	CpG island	CpG island shore	Other region	Total
32	188	734	10803	4250	2490	5017	11757

- Percent methylation difference > 25%
- DE genes



Results – pathway

39 pathways enriched by 13 genes included *EGFR*, *PEMT*

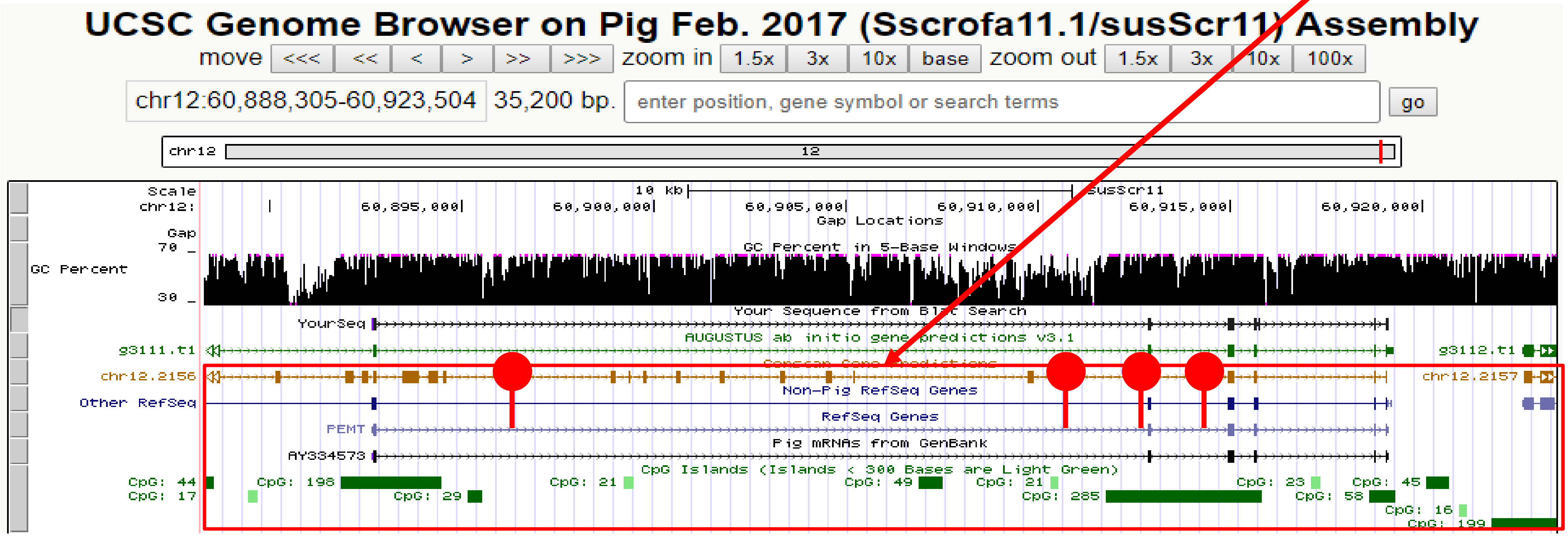


Results – candidate biomarker

DMCs located within DE genes

Chromosome	DE gene	Gene description	Structure	DMC
6	*DMAP1*	DNA methyltransferase 1 associated protein 1	Up 10K upstream	${}^{a2}2$
9	*EGFR*	Epidermal growth factor receptor	Gene body	${}^{a2b1}3$
12	*PEMT*	Phosphatidylethanolamine N-methyltransferase	Gene body	${}^{a2b2}9$

Note: $^{a/b}$Number of DMCs located within CpG islands/CpG island shores.


UCSC Genome Browser on Pig Feb. 2017 (Sscrofa11.1/susScr11) Assembly

chr12:60,888,305-60,923,504 35,200 bp.

Discussion

- **This is first study to report Genome-wide DNA methylation profiles of BT trait in pig using NGS methods**

- **Our results evaluated candidate genes for example *DMAP1*, *EGFR* and *PEMT*, associated with epigenetic DMCs from pig genome**

Acknowledgments

- **This study was funded by the AGES project and the GUDP project (PI: Haja Kadarmideen) - both projects received funding from Danish Ministry of Food, Agriculture and Fisheries.**
- **Xiao Wang received Ph.D. stipends from the Technical University of Denmark, and the China Scholarship Council, China.**
- **We thank SEGES-Pig Research Center (VSP) for collaboration in all of our previous boar taint projects**

Acknowledgments

- Prof. Haja N. Kadarmideen, main PI / group leader who conceived and designed all boar taint experiments and supervision of this work

- Markus Drag and Dr. Ruta Skinkte, who produced the RRBS data, provided scientific support on boar taint

- Dr. Gianluca Mazzoni, who provided scientific support on organization of results

- Members of QSG / QGBC group at DTU Bioinformatics & DTU Compute

Thank you for the attention!