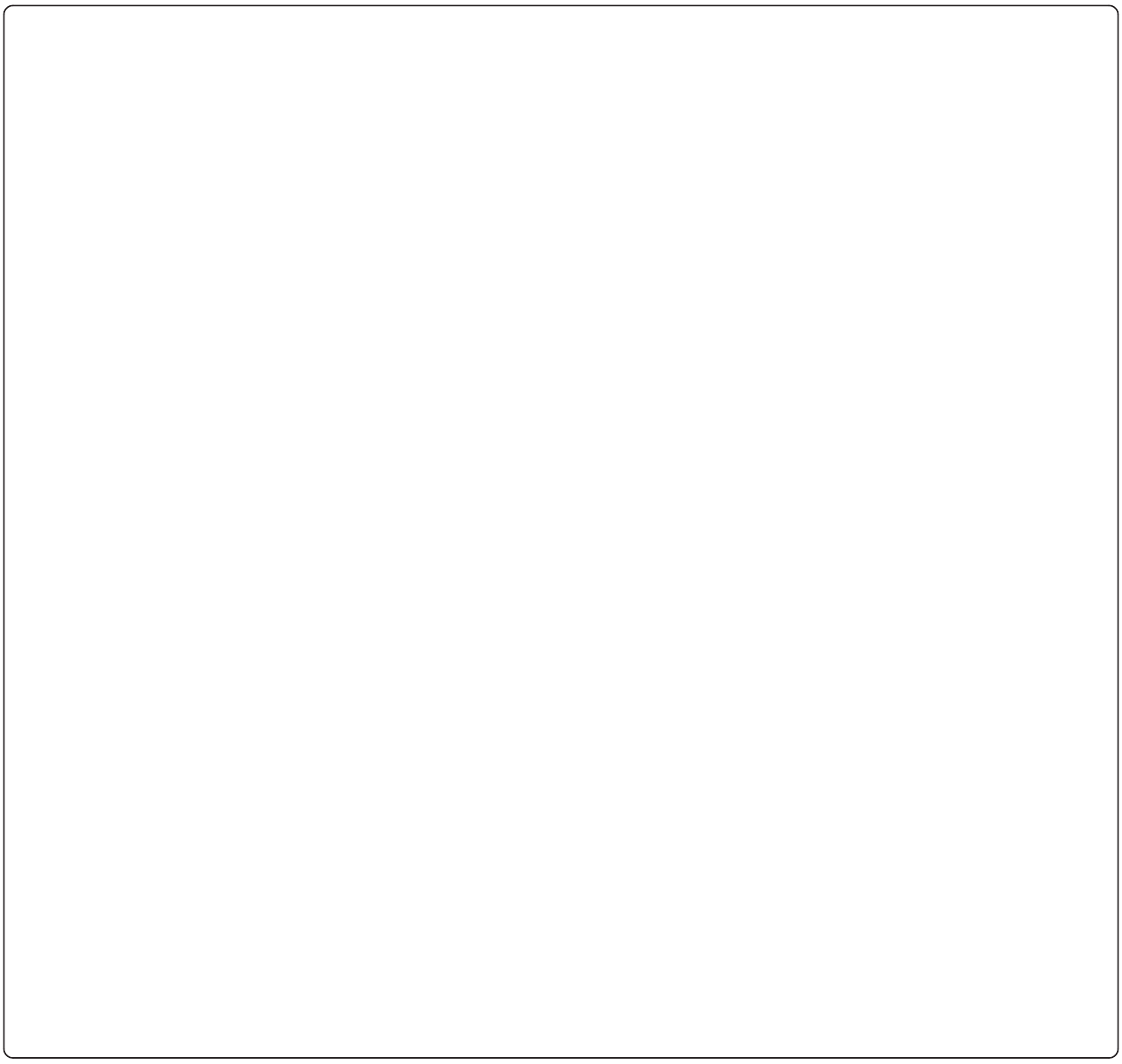
transposable elements. Integration with our transcriptome assembly reveals that transcription correlates accurately with DNA methylation and accounts for approximately 85–90 % of the methylome. We generate a mouse model in which transcription across the locus is abrogated in oocytes, resulting in failure of DNA methylation establishment at all CpGs of this locus. ChIP analysis in oocytes reveals H3K4me2 enrichment at the imprinted control region when transcription is ablated, establishing a connection between transcription and chromatin remodeling at CpG islands by histone demethylases.

Conclusions: By precisely defining the mouse oocyte transcriptome, this work not only highlights transcription as a



F 1 A 1), I . - (F 2 A 1). , - C (14) 100 % - C C (C C) C - - (F 2 A 1). -C I (C CI) 28, 29 . C C 841 23.1 % - 834 - (18.6 % , , ,) C CI 100 - 188 - (2.5 (3.2 %) . E - 19.1 ,) 7.6 % - . I (C C C CI , - — 171), (F A . I , 2.65 0.19, , G 2 ; F . 2). -

( ). A
D A ,

C G . , . 52
E- 2.9 % (41 -
5.7 % CGI , .) ,
-
Some expressed genes escape DNA methylation (3 A 2). ,
I - -
D A , , D A . A -
D A . 41 ,
, H D - -
. 26.2 % H D (
(>50 %), 16.1 %
H D , (F . 3 ; F 5 A 1).
, H D (9.4), 51.9 % 23.3 %
10 5 , . -
-
H3 36 3 37 .
H D
F ,
,
D A
,
H3 36 3 4 ; , <0.5
F ,
F (F 5 A 1).
46.2 % H D (14.7 , 11.0 %
H D) <0.5 F .
I , H D
(14.2 %; 972) -
; (4.4) -
H3 4 2/ 3 15 (F . 3 ; F 5
A 1).
- D A -
, 9.2 % H D (3.7 %
H D) (>50 %)
(F 5 A 1).
D A . 318
- D A <25 %,
(F > 1 10
). ,
( ) -
-
.
C ,

(C - ı , p ı <0.0001). ı , -
ı D A 3′
ı (40 % ı
, D A ı 75 % ı 1 -
ı) ı 18.7 % ı CGI -
ı 1 ı ı . F
ı ı CGI , ı
ı G ,
D 3A ı D 3 , ı ,
ı ı ı .
ı ı CGI ı ı -
ı D A , ı
ı ı - ı . 2863 ı -
ı CGI , 41.5 % ı
(2) ı , ı -
, ı “ ı ”

Zac1o -
Sox2-Cre
D ,
(Zac1) (F 7 A 1).
Zac1 D -
, -
Zac1 -
, Zac1o−/−
Mus castaneus (F . 4). -
,
Zac1 D
C I Zac1o+/+ Zac1o+/− (-
) (F 7 A 1).
Zac1o -
, D
. ,
D -
: H3 4 3 H3 9
, ,

Zac1o+/+ ,
H3 9 3
H4 20 3, -
Zac1o+/+ .
9.5
Dnmt3L−/− -
D A 42 .

Transcription is required for full chromatin remodelling at the ***Zac1*** igDMR

Zac1o ,
D A . A
H3 36 3 D 3A , -
CGI , 2 . I -
, H3 4 2
D 1B D A
Zac1 D 43 , D 1B . F
C I - C

-C I 44 . -
2000 (15)
(Zac1o+/+ Zac1o–/–),
-
C
(, , D).

G 1 G 2 .2
A- (E).
G FG -
III (-
), D A
D A I (EB); -
EB D A
I (EB),
E E (EB) C .

Library sequencing and mapping

G , G 1, G 2 FG A-
100- - I

E , C C
C C B 1
2014. G C
C .2.1.1 .
-
.
C , =, , ,
, ,
.
CGI D 7, 9, 49,
50 C C
G C 38 , CGI .
CGI
100 ,
, -
. C E (1
2 I E , 2 4 I E , E 1 , E , E ,
) G C 38
.
CGI- 100- CGI E-
E E .
E ,

A a 2: Supplementary Tables S1–S4. (PDF 240 kb)

A a 3: Supplementary methods, including more details on the transcriptome assembly process. (PDF 107 kb)

A a 4: A table with the genomic coordinates of our final definition of HyperDs and HypoDs (GRCm38 assembly). (XLSX 4255 kb)

A a 5: A gtf file corresponding to the annotation of oocyte transcripts according to our transcriptome assembly (GRCm38 assembly; this version does not include expressed independent repetitive elements). (GTF 50970 kb)

Abbreviations

bp: base pair; BS: bisulfite sequencing; CGI: CpG island; ChIP: chromatin immunoprecipitation; CNCI: Coding-Non-Coding Index; CPC: Coding Potential Calculator; DNAme: DNA methylation; DNMT: DNA methyltransferase; dpp: days post-partum; E: embryonic day; ESC: embryonic stem cell; FGO: fully grown oocyte; FPKM: fragments per kilobase of transcript per million mapped reads; GO: growing oocyte; HyperD: hypermethylated domain; HypoD: hypomethylated domain; igDMR: imprinted germline differentially methylated regions; ncRNA: non-coding RNA; NGO: non-growing oocyte; PCR: polymerase chain reaction; PGC: primordial germ cell; RABT: reference annotation-based transcript; RNA-Seq: RNA sequencing; RRBS: reduced representation bisulfite sequencing; TE: transposable element; TSS: transcription start site; UCSC: University of California, Santa Cruz.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LV performed RNA-Seq, transcriptome assembly, and global correlation with transcription. SAS. generated conditional knockout mice, and performed -related experiments except ChIP in embryos, which was performed by SM-M, and PA, LV, HS, FK, and SA performed bioinformatic analyses. KS and ST provided material and advice. GK and SAS initiated and supervised the project. SAS wrote the manuscript with input from LV and GK. All authors read and approved the final manuscript.

Acknowledgements

We thank all members of the Kelsey laboratory and Wendy Dean for helpful discussions, Kristina Tabbada for assistance with Illumina sequencing, and Solenn Patalano for help with RNA-Seq methodology. This work was supported by MRC Programme grant MR/K011332/1, MRC Project grant G0800013, FP6 SAVEBETA, and a Babraham Institute and University of Cambridge graduate student scholarship.

Author details

[1]Epigenetics Programme, Babraham Institute, Cambridge, UK. [2]Bioinformatics Group, Babraham Institute, Cambridge, UK. [3]GReD, CNRS, INSERM, and Clermont University, 63001 Clermont-Ferrand, France. [4]Department of Histology and Cell Biology, Yokohama City University School of Medicine, Yokohama, Japan. [5]Centre for Trophoblast Research, University of Cambridge, Cambridge, UK.

R c : 24 J- 2015 Acc : 3 S b 2015