

SHORT COMMUNICATION

# Systems genomics analysis centered on epigenetic inheritance supports development of a unified theory of biology

Abhay Sharma\*

## ABSTRACT

New discoveries are increasingly demanding integration of epigenetics, molecular biology, genomic networks and physiology with evolution. This article provides a proof of concept for evolutionary transgenerational systems biology, proposed recently in the context of epigenetic inheritance in mammals. Gene set enrichment analysis of available genome-level mammalian data presented here seem consistent with the concept that: (1) heritable information about environmental effects in somatic cells is communicated to the germline by circulating microRNAs (miRNAs) or other RNAs released in physiological fluids; (2) epigenetic factors including miRNA-like small RNAs, DNA methylation and histone modifications are propagated across generations via gene networks; and (3) inherited epigenetic variations in the form of methylated cytosines are fixed in the population as thymines over the evolutionary time course. The analysis supports integration of physiology and epigenetics with inheritance and evolution. This may catalyze efforts to develop a unified theory of biology.

**KEY WORDS:** MicroRNA, Gene expression and networks, DNA methylation, Histone modification

## INTRODUCTION

Because of the inability of contemporary gene and natural selection centric evolutionary theory to fully explain heritability of phenotypic traits based solely on DNA sequence variation on the one hand, and increasing evidence of non-genetic inheritance on the other, there is a profound interest in integrating epigenetics, molecular biology, genomic networks, and physiology with the theory of evolution (Petronis, 2010; Danchin et al., 2011; Mattick, 2012; Ball, 2013; Noble, 2013, 2015; Noble et al., 2014). A top-down approach toward this unification may begin with developing a broad conceptual mechanistic framework and testing that in a proof-of-concept analysis using empirical data. Notably, a recently proposed framework that is supported by observations reported in the literature (Sharma, 2014, 2015a,b,c) provides an opportunity to analytically test the concept of a unified theory of biology. The proposed model explains transgenerational epigenetic inheritance and its evolutionary significance by integrating gene expression and gene networks, miRNA or other RNA, DNA methylation, histone modifications, and DNA-methylation-induced mutation, in three mechanistic steps (Fig. S1). First, heritable information about environmental effects in somatic cells is communicated to the germline by circulating RNAs, representing physiological conditions. Second, epigenetic modifications, in the form of

RNA, DNA methylation and histone modifications, are propagated across generations through gene expression and gene networks. Third, inherited epigenetic variations, represented by methylated cytosines, are fixed in the population as thymines, in an evolutionary time scale. In the present analysis, these three principles are tested using available large-scale data sets related to gene expression, DNA methylation, histone modification, cytosine-methylation-induced mutation, chemical gene interaction and genetic association in diseases. The analysis is based on gene set enrichment as a measure of association. Given that the original framework was proposed in the context of mammals (Sharma, 2014, 2015a,b,c), only mammalian data is used here for model validation.

## RESULTS AND DISCUSSION

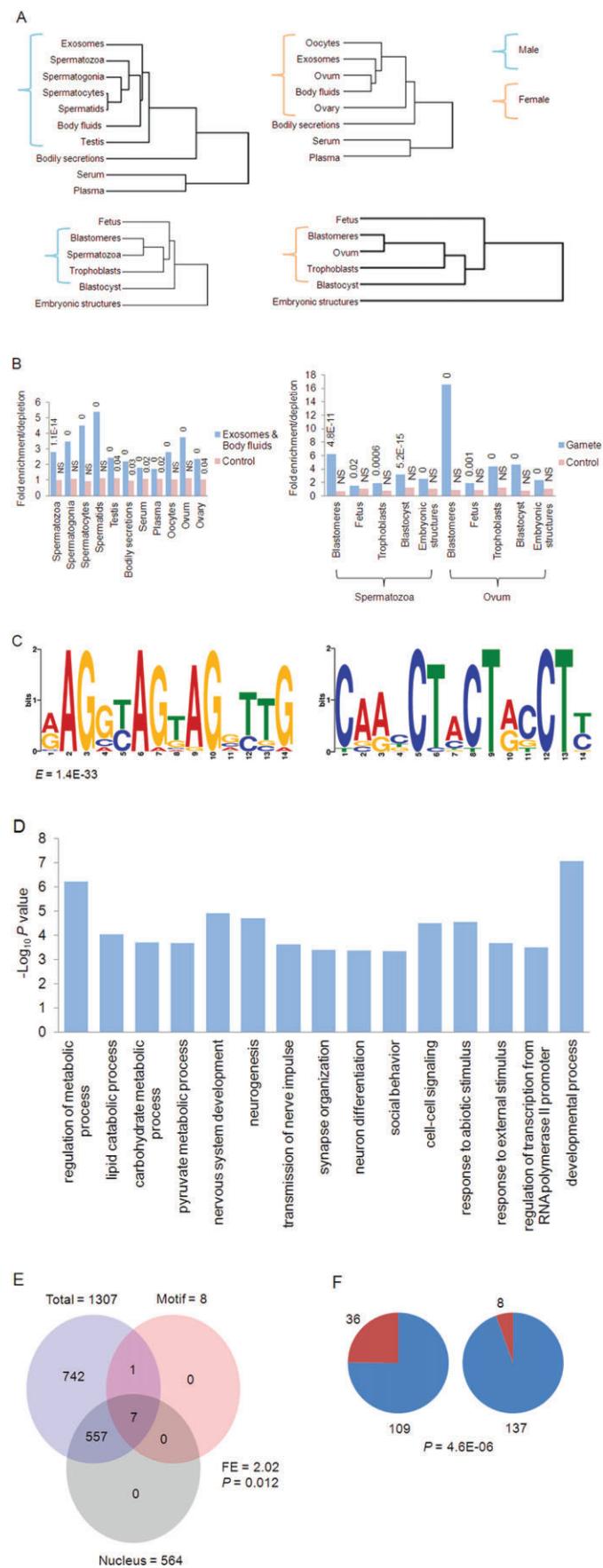
Fig. 1 illustrates evidence supporting the first principle. The circulating miRNA profiles of exosomes and body fluids closely resemble that of gonads and germ cells, in both males and females, and the miRNA profiles of germ cells resemble that of embryonic stages (Fig. 1A). The miRNA-based similarity observed between circulating factors, germ cells and developing embryo is statistically significant (Fig. 1B). miRNAs of exosome and body fluids were found to be over-represented in germ cells, as were miRNAs of gametes in the developing embryo. In contrast, number-matched control miRNAs, selected randomly from the combined set of miRNAs representing 461 organs, tissues and various other samples in the database, were not over-represented in general. These results clearly suggest that gametogenesis and development are strongly related to circulating factors in terms of miRNA profile.

Secreted by all cell types and identified in diverse body fluids, exosomes carry miRNAs that can influence gene expression and cause physiological changes in recipient cells (Chevillet et al., 2014; Melo et al., 2014; Alexander et al., 2015). Exosomal miRNAs are also considered to play a role in male and female reproductive physiology in mammals (Belleanné et al., 2013; Santonocito et al., 2014; Barkalina et al., 2015). Evidence suggests that miRNAs are loaded into exosomes selectively, based on specific miRNA motifs and post-transcriptional modifications, and levels of miRNAs or their targets in the producer cells (Alexander et al., 2015). The preferential sorting is supported by the observation that miRNA signatures of exosomes do not directly mirror the miRNA composition of the producer cells (Alexander et al., 2015). These findings have suggested that some miRNAs may have evolved to be parceled out in exosomes to perform their biological roles (Alexander et al., 2015). Although separating exosomal miRNAs from other extracellular-vesicle-associated or vesicle-independent circulating miRNAs is technically challenging, it has recently been inferred from quantitative and stoichiometric analysis that most exosomes do not contain many copies of miRNA molecules (Chevillet et al., 2014). This has led to the hypothesis that a given miRNA is either distributed in a few exosomes in the population at low concentration or, alternatively, packaged in rare exosomes at

CSIR-Institute of Genomics and Integrative Biology, Council of Scientific and Industrial Research, Sukhdev Vihar, Mathura Road, New Delhi 110025, India.

\*Author for correspondence (abhaysharma@igib.res.in)

Received 26 May 2015; Accepted 20 August 2015

**Fig. 1. Soma-to-germline communication of phenotypic information.**

(A) Clustering of biological samples based on presence or absence of microRNAs (miRNAs). Exosomal and body fluid profiles match closely with that of gonads and gametes, and the profiles of gametes with that of developmental stages. (B) Over-representation of exosomal and body fluid miRNAs combined in gonads and gametes, and of gametic miRNAs in developmental stages. (C) Display of the motif discovered. The motif logo, with reverse complement on the right, is shown along with the *E* value. (D) A subset of enriched gene ontology biological processes, all with *q*-value<0.05, in genes with motif-containing promoters. Log-transformed nominal *P*-values for individual enrichment are plotted on the *y*-axis. (E) Venn diagram showing enrichment of the motif-containing miRNAs in the set of nucleus-localized miRNAs. Fold enrichment (FE) and enrichment *P*-value are indicated. (F) Pie chart depicting enrichment of the motif in the set of nucleus-localized miRNAs excluding the initially discovered motif-containing miRNAs. Enrichment *P*-values, hypergeometric distribution, related to data pertaining to bar diagrams are shown above the bars. Values with 16 or more negative exponents of 10 were rounded down to zero. NS, not significant. List of miRNAs representing all the samples indicated in A,B are given in Table S1. Details of motif discovery analysis in C are provided in Table S2. Complete set of enriched gene ontology biological processes in D is shown in Table S3. Details of motif enrichment in F are given in Table S4.

high concentration (Chevillet et al., 2014). Furthermore, selective loss of certain miRNAs with low GC content during extraction from smaller samples has been noted (Kim et al., 2012). Given these mechanistic and technical reasons associated with quantitative analysis of extracellular miRNAs, the present qualitative analysis, based simply on the presence or absence of a given miRNA, showing profile similarity between a general pool of circulating miRNAs, and gonadal and gamete-borne miRNAs may at this time seem sufficient to support the concept that extracellular noncoding RNA can potentially mediate soma to germline transmission of heritable information.

To examine the possibility that certain circulating miRNAs may have evolved to be selectively released into the circulation to mediate epigenetic inheritance, the exosome and body fluid miRNAs present in testis, spermatogonia, spermatocytes, spermatids or spermatozoa, as well as in the ovary, oocytes or ovum were examined for *de novo* discovery of sequence motifs. Interestingly, a motif was discovered in these miRNAs, not in a number-matched control set of randomly selected exosome and body fluid miRNAs (Fig. 1C). The motif was present in 19 of the total 131 miRNAs. Startlingly, enrichment analysis showed that the motif was significantly over-represented in promoters, 1000 base pair upstream to 200 base pair downstream regions, of genes associated with various gene ontology terms, including those that seem consistent with epigenetic inheritance, related to environmental factors, intercellular communication, gene expression, development, energy metabolism, and nervous system structure and function, for example (Fig. 1D). It is notable that a large proportion of examples of non-genetic inheritance reported so far in mammals relate to carbohydrate and lipid metabolism, and brain and behavior (Choi and Mango, 2014; Bale, 2015; Sharma, 2015b; Szyf, 2015; Dias et al., 2015). In addition, a recent study has found that thousands of genes that escape genome-wide DNA demethylation in human primordial germ cells (PGCs), whose complete data set is not reported, are over-represented in genes expressed in brain, and genes associated with metabolic, and neurological and neuropsychiatric disorders (Tang et al., 2015). Importantly, the potential biological significance of the identified miRNA motif and its presence in promoters was supported by examining nuclear localization of the miRNAs. The raw nuclear

read counts for 8 of the 19 motif-containing miRNAs were available in a reported set of human cell line small RNA deep sequencing data, wherein a count of 10 or more was considered to indicate nuclear localization (Liao et al., 2010). Interestingly, in the data spanning 1307 unique mature miRNA sequences, the motif-containing miRNAs were over-represented in nucleus-localized miRNAs (Fig. 1E), with 7 of the 8 motif-containing miRNAs figuring within the top 21 miRNAs with highest counts in the nucleus. Furthermore, the top 145 nucleus-localized miRNAs, arbitrarily chosen from 1307 sequences, excluding the above 7 miRNAs, were found to be highly enriched for the motif, compared with the bottom 145 sequences with nuclear counts below the threshold of that for nuclear localization (Fig. 1F). Thus, profound nuclear localization of the motif-containing miRNAs and prominent presence of the motif in promoters of certain categories of genes together support the possibility that these miRNAs may regulate gene expression at transcriptional level. Indeed, examples for miRNA-mediated transcriptional regulation are known (Zhang et al., 2014). In particular, evidence exists to suggest that miRNAs with binding sites in gene promoters, located within ~1000 base pairs without any unique feature, can modulate gene expression through epigenetic modifications of the promoter, including histone acetylation and/or methylation (Zhang et al., 2014). This is consistent with the present analysis supporting a role of circulating miRNAs in epigenetic inheritance.

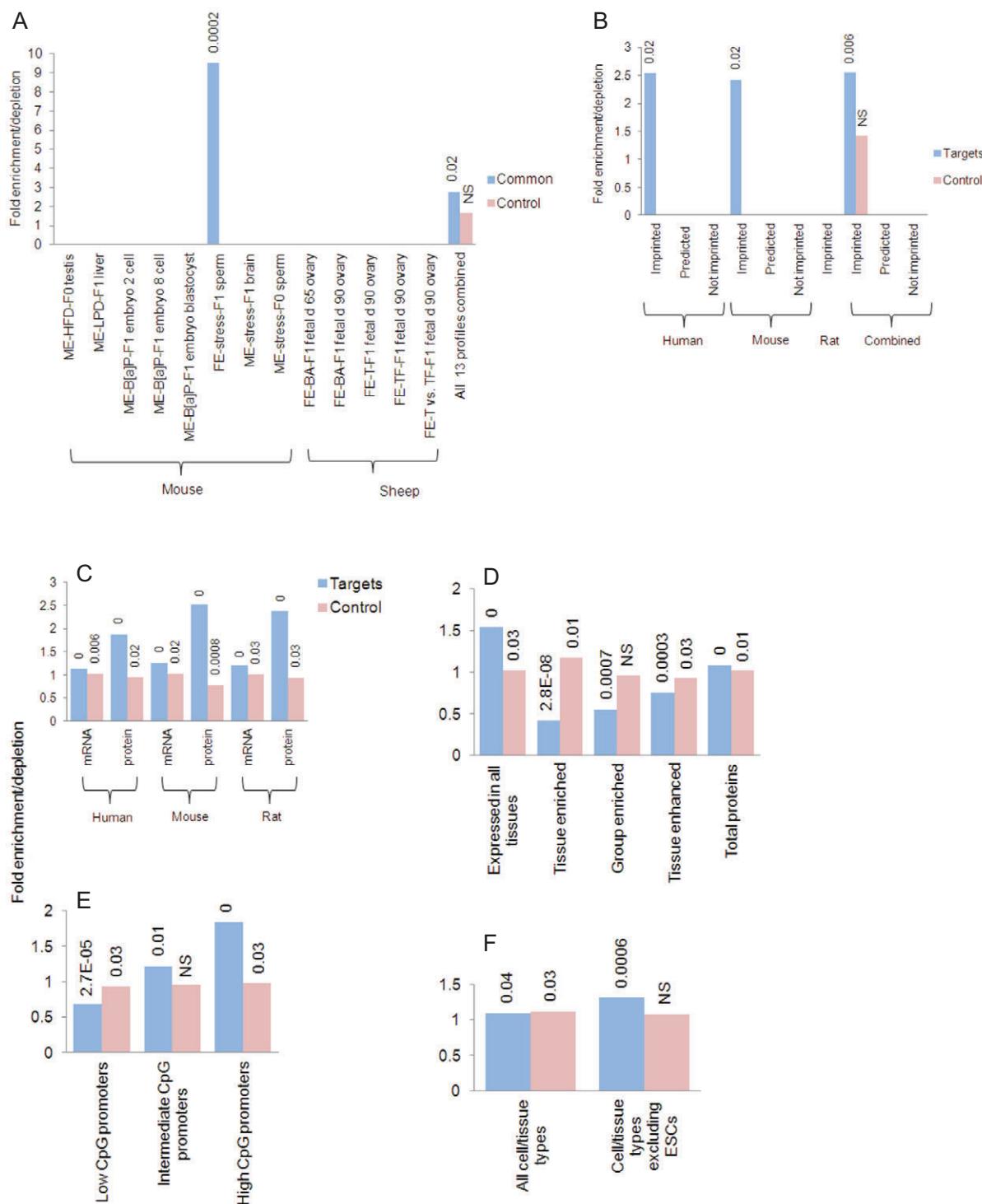
A mechanism of multigenerational epigenetic inheritance mediated by Piwi-interacting RNA (piRNA), a class of small noncoding RNAs that are expressed in male and female germline and play an evolutionarily conserved role in transposon silencing, has previously been demonstrated in the nematode *Caenorhabditis elegans* (Ashe et al., 2012). In that study, it was shown for the first time that a piRNA-dependent foreign RNA response leads to multigenerational gene silencing involving a germline nuclear small RNA/chromatin pathway. In *C. elegans*, the mechanisms underlying piRNA mediated transcriptional gene silencing are considered similar to that involved in nuclear RNAi pathway in somatic tissues, with repressive histone modifications and RNA polymerase II stalling leading to silencing (Weick and Miska, 2014). The present analysis raising the possibility of miRNA, an endogenous small RNA like piRNA, playing a role in epigenetic inheritance in mammals seems very attractive because it satisfies the requirement of soma to germline communication, as envisaged in the first principle of the model under investigation.

Fig. 2 and Figs S2–S7 display evidence in support of the second principle. The common circulating and germline miRNAs are over-represented among miRNAs identified as differentially expressed in studies examining environmental effects in exposed and unexposed generations (Fig. 2A). In addition, the target genes of these common miRNAs show enrichment for genes that show differential mRNA expression (Fig. S2) and DNA methylation (Fig. S3) in gonads, gametes and various other tissues and organs in transgenerational studies. These results suggest a role of gene networks in epigenetic inheritance. Global analysis of data representing normal conditions further supports this. For example, the targets are found to over-represent imprinted genes (Fig. 2B), with imprinting representing a mode of epigenetic inheritance. The targets also enrich genes that are known to interact with a broad range of chemicals (Fig. 2C), including environmental factors known to cause transgenerational effects. Genes showing expression, regulation and differential DNA methylation, and histone modifications in gametes and developing embryo under normal conditions are also over-represented in the targets (Figs S4–S6). Furthermore, the targets enrich genes showing

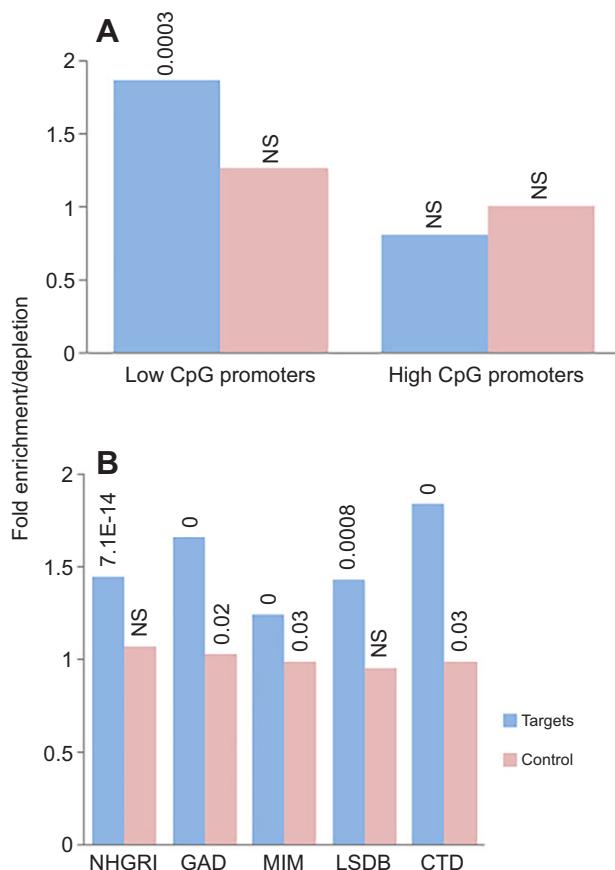
tissue-wide expression (Fig. 2D), differential genome-level CpG promoter density distribution (Fig. 2E) and histone modification signatures of active promoters and enhancers (Fig. 2F). The targets are also found to over-represent genes with known function in processes related to gene regulation by noncoding RNA including miRNA, biogenesis and metabolism of these RNAs, gene-specific transcription, response to chemicals and abiotic factors, embryonic development (Fig. S7) and metabolism, and brain development and function (Table S7). Finally, the reported binding regions of BLIMP1 (B lymphocyte-induced maturation protein-1; Magnúsdóttir et al., 2013), a key regulator of PGC specification involved in resetting of the epigenome towards a basal state, were found to be highly significantly enriched in the targets (fold change, 1.7;  $P < 1 \times 10^{-16}$ ), with the control showing a slight depletion (fold change, 0.86;  $P < 0.01$ ). Cumulatively, these results appear consistent with the second principle of the conceptual framework implicating gene networks in epigenetic inheritance.

Fig. 3 presents evidence supporting the third principle. The targets show enrichment for transcription factor binding sites created by cytosine-methylation-induced mutation in low-CpG promoter-associated genes (Fig. 3A). As CpG dinucleotides in these promoters are constitutively methylated in the germline (Weber et al., 2007), this result supports the evolutionary significance of epigenetic inheritance. Next, the targets were found to over-represent genes mutations or polymorphisms that show an association with various diseases (Fig. 3B), potentially connecting epigenetic variations, evolution and human health. This is consistent with the above-mentioned study showing over-representation of disease-associated genes in DNA-demethylation-resistant genes in human PGCs (Tang et al., 2015). As proposed (Sharma, 2015b,c), the epigenetic modifications may either discontinue to exist, persist as such or convert to genetic alterations in evolutionary time course (Fig. S1). Although transition of 5-methylcytosines to thymines per se could be a passive by-product, the resulting change can become a potential substrate for selection like any other newly arisen genetic variation does in the normal course of evolution. A recently proposed model of RNA-mediated gene evolution has underscored this possibility (Morris, 2015). With regards the overall evolutionary significance of epigenetic inheritance, it has been highlighted in a recently advanced unified theory of molecular aspects of evolution that natural selection may potentially act on acquired traits (Skinner, 2015). Cumulatively, the present results seem consistent with the third principle.

Evidence provided here supports the proposed (Sharma, 2014, 2015a,b,c) model of evolutionary transgenerational systems biology. First, it demonstrates that miRNA-based similarity exists between circulating factors, gametes, developmental stages and adult tissues, under normal conditions or under environmental conditions that produce an effect across generations. This is consistent with the hypothesis that RNAs present in physiological fluids hold the potential to mediate soma to germline communication in epigenetic inheritance. Second, the results show an association between the target genes of the convergent miRNAs and regulated gene expression, DNA methylation and histone modifications in the above conditions. This is in line with the proposition that gene networks may underlie epigenetic memory propagation across generations. Third, the miRNA targets show association with cytosine-methylation-induced mutational events and disease-related mutations and polymorphisms. This supports the suggestion that epigenetic inheritance may play an evolutionarily significant role. The above analysis is not exhaustive in the sense that



**Fig. 2. Gene networks in germline transmission.** (A) Over-representation of miRNAs common to exosomes, body fluids, spermatozoa, ovum and all other tissues indicated below the bars shown in Fig. 1B except fetus, total 18 in number, in the sets representing differentially expressed miRNAs reported in various studies on epigenetic inheritance. Note combined analysis, represented by the extreme right pair of bars, showing enrichment of common miRNAs in total miRNAs reported in epigenetic studies. Key to miRNA profiles: gender exposed at F0-treatment at F0-generation and sample profiled. ME, male exposure; FE, female exposure; HFD, high-fat diet; LPD, low-protein diet; T, testosterone; F, flutamide; TF, testosterone plus flutamide; B[a]P, benzo(a)pyrene; BA, bisphenol A. Mature miRNA names were used for enrichment analysis. (B) Enrichment of targets (see Fig. S2) in imprinted genes. (C) Enrichment of targets (see Fig. S2) in chemical-interacting mRNAs and proteins. (D) Proteomic profiles across human tissues. Over-representation of the targets in proteins expressed in all tissues, and depletion in proteins with tissue biased expression, are clearly observed. (E) Genome-level CpG density across human promoters. The targets are enriched and depleted in high- and low-CpG promoters, in that order. (F) Histone modification signatures of active promoters and enhancers in human cell/tissue types. The targets over-represent active regulatory signatures in cells/tissues other than embryonic stem cells (ESCs). Enrichment P-values, hypergeometric distribution, are shown above bars. Values with 16 or more negative exponents of 10 were rounded down to zero. NS, not significant. Missing bars in A,B indicate overlapping counts as either nil or excessively low, considered inappropriate for analysis. Sources of data used in all analyses are indicated in Table S5. Complete set of enriched gene ontology biological processes is provided in Table S6.



**Fig. 3. Evolutionary significance of epigenetic inheritance.**

(A) Representation of the targets in gene promoters associated with transcription factor binding sites created by mutation of cytosine to thymine in the context of CpG dinucleotides. Targets are clearly enriched in low-CpG promoter-associated genes. (B) Enrichment of the targets in disease-associated genes. Targets are clearly enriched. NHGRI, National Human Genome Research Institute; GAD, genetic association database; MIM, Mendelian inheritance in man; LSDB, locus-specific mutation databases; CTD, comparative toxicogenomics database. Enrichment *P*-values, hypergeometric distribution, are shown above bars. NS, not significant. Values with 16 or more negative exponents of 10 were rounded down to zero. Sources of data used in the analysis are indicated in Table S7.

several sets of available data remain to be examined. For example, a recent study has revealed that several thousand genes escape genome-wide DNA demethylation in human PGCs (Tang et al., 2015). Non-availability of the complete data set, however, prevented its analysis here. But this study and several others (Magnúsdóttir et al., 2013; Smith et al., 2014; Irie et al., 2015) together provide additional data on gene expression and DNA methylation dynamics associated with gametes, and cells, tissues and organs pertaining to embryonic development in mouse and human. It will be interesting to extend the analysis to include these and newer studies in future. Nevertheless, the present results tend to unify environment, physiology, RNA, gene networks, DNA methylation and histone modifications with inheritance, development, disease and evolution. A potential role and integration of physiology, epigenetics, RNA and genetic interactions in inheritance and evolution have previously been suggested and sought for on the basis of theoretical considerations (Richards, 2008; Day and Bonduriansky, 2011; Hunter et al., 2012; Livnat, 2013; Noble, 2013, 2015; Noble et al., 2014; Rivoire and Leibler, 2014; Jablonka and Lamb, 2015; Morris, 2015; Skinner, 2015). The data analysis presented here offers a proof

of principle for a unified theory of biology. The results may prove valuable in directing future efforts toward integration.

## MATERIALS AND METHODS

Data available in various public databases and published papers along with associated supplementary information was used, as provided. The papers were identified through PubMed search using appropriate key words, and the relevant data sets included in the analysis without any bias. For miRNA and mRNA clustering or enrichment analysis, all the human miRNAs in the organ-miRNA interaction database (<http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk>) or all the human genes in gene ontology (<http://geneontology.org/>) were used as the total miRNA or mRNA population, in that order. To obtain controls for enrichment analysis, matching numbers of miRNAs or mRNAs were randomly selected (<https://www.random.org/sequences/>) from the aforesaid populations, as appropriate. A documented set of validated miRNA target genes was used (<http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/index.html>). The statistical significance of enrichment was computed using hypergeometric distribution probability, with 0.05 as the nominal *P*-value cut-off. Over-representation analysis of gene ontology biological processes was carried out using 0.05 as Benjamini-Hochberg adjusted *P*-value cut-off (<http://david.abcc.ncifcrf.gov/>). Chi-square test for homogeneity was used to check equality of two different populations. The MEME suite 4.10.1 (Bailey et al., 2009) was used for *de novo* discovery ([meme-suite.org/tools/meme](http://meme-suite.org/tools/meme)) and enrichment ([meme-suite.org/tools/fimo](http://meme-suite.org/tools/fimo)) of miRNA motif, and its association with promoters of genes linked to gene ontology terms ([meme-suite.org/tools/gomo](http://meme-suite.org/tools/gomo)), all under default settings. For motif discovery, the complete set of human miRNAs in miRBase (<http://www.mirbase.org/>) was used as background (Kozomara and Griffiths-Jones, 2014). A 0-order Markov model was assumed for the background, and the minimum motif width set at 6 nucleotides.

## Competing interests

The author declares no competing or financial interests.

## Funding

This research received a grant from the BSC0122 network project of the Council of Scientific and Industrial Research, India.

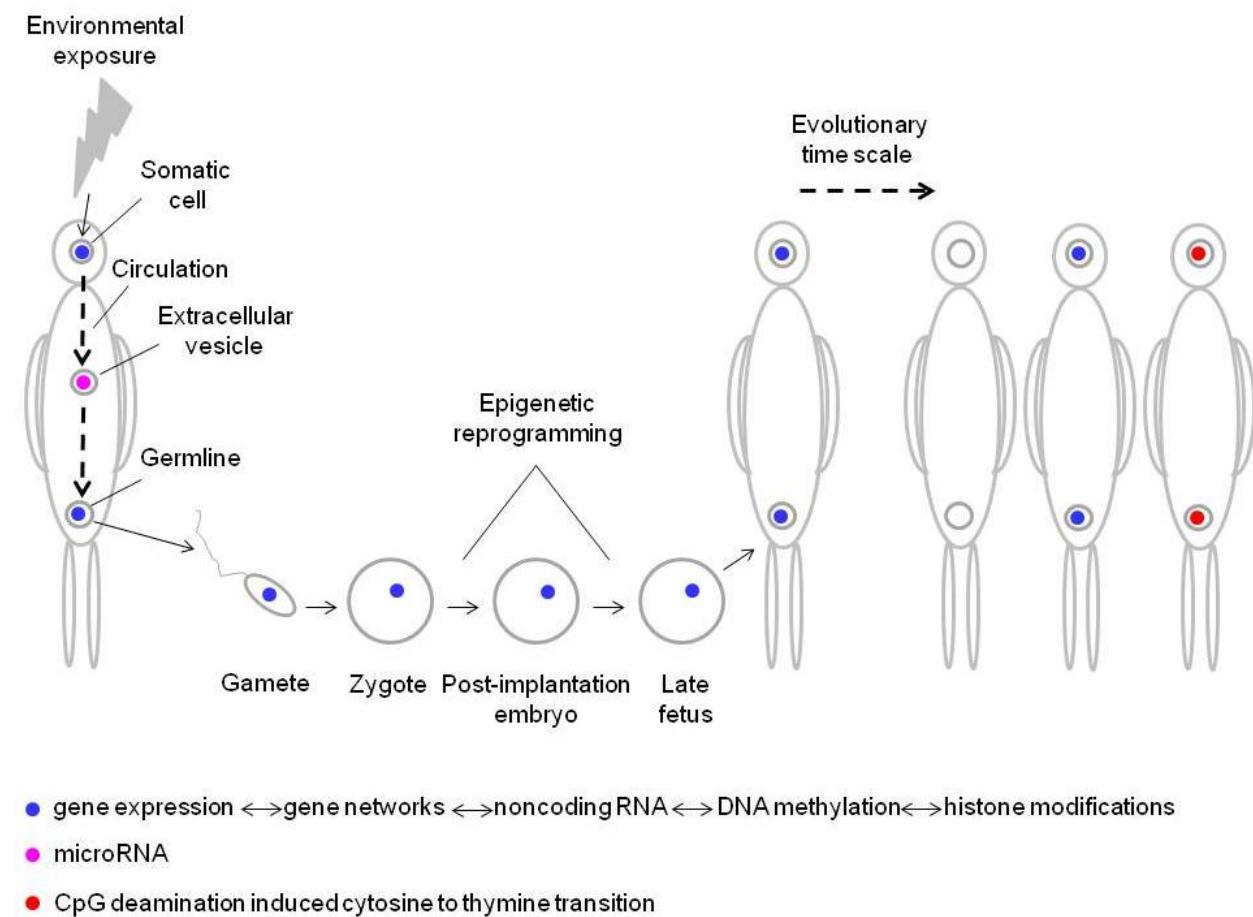
## Supplementary information

Supplementary information available online at <http://jeb.biologists.org/lookup/suppl/doi:10.1242/jeb.125922/-DC1>

## References

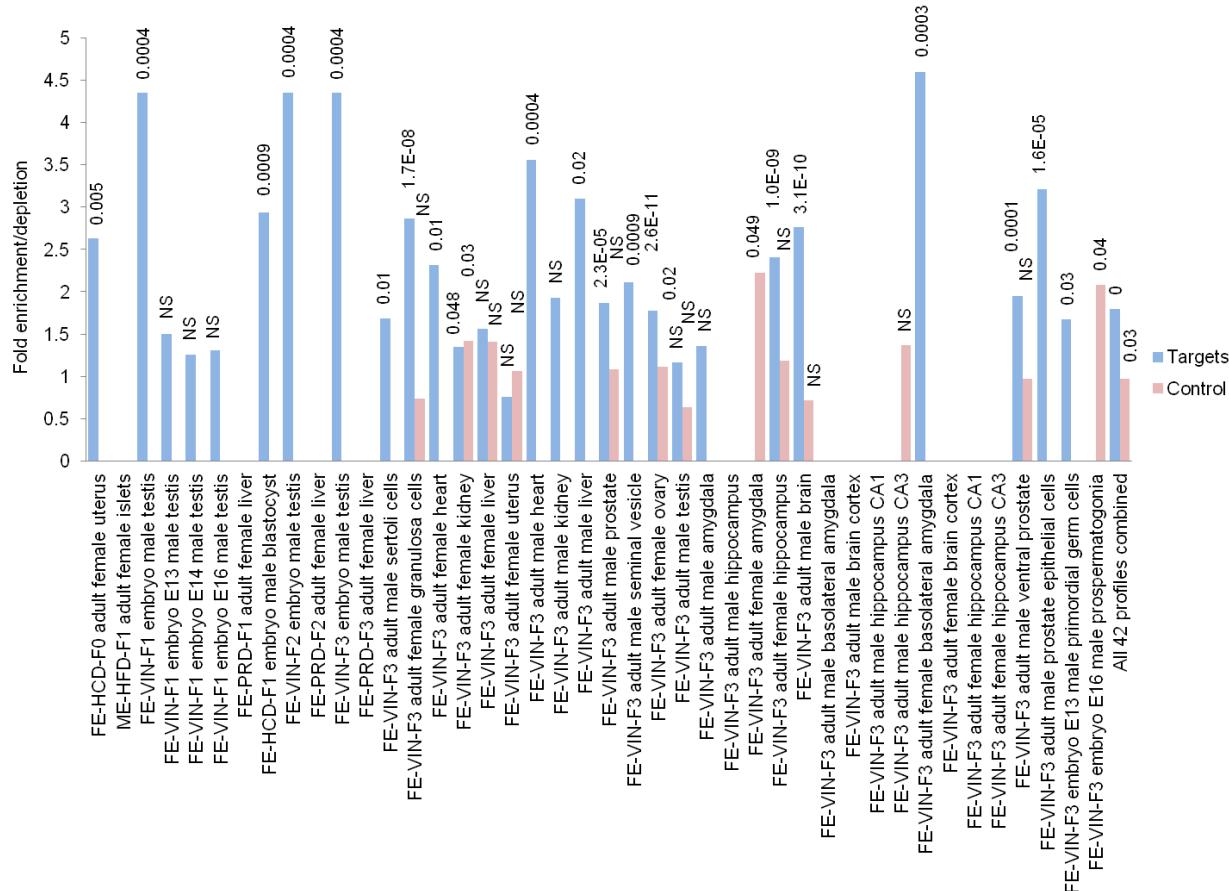
- Alexander, M., Hu, R., Runtsch, M. C., Kagele, D. A., Mosbruger, T. L., Tolmachova, T., Seabra, M. C., Round, J. L., Ward, D. M. and O'Connell, R. M. (2015). Exosome-delivered microRNAs modulate the inflammatory response to endotoxin. *Nat. Commun.* **6**, 7321.
- Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M. P., Cording, A. C., Doebley, A.-L., Goldstein, L. D., Lehrbach, N. J., Le Pen, J. et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* **150**, 88-99.
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. and Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202-W208.
- Bale, T. L. (2015). Epigenetic and transgenerational reprogramming of brain development. *Nat. Rev. Neurosci.* **16**, 332-344.
- Ball, P. (2013). DNA: celebrate the unknowns. *Nature* **496**, 419-420.
- Barkalina, N., Jones, C., Wood, M. J. A. and Coward, K. (2015). Extracellular vesicle-mediated delivery of molecular compounds into gametes and embryos: learning from nature. *Hum. Reprod. Update* **21**, 627-639.
- Belleanné, C., Calvo, É., Caballero, J. and Sullivan, R. (2013). Epididymosomes convey different repertoires of microRNAs throughout the bovine epididymis. *Biol. Reprod.* **89**, 30.
- Chevillet, J. R., Kang, Q., Ruf, I. K., Briggs, H. A., Vojtech, L. N., Hughes, S. M., Cheng, H. H., Arroyo, J. D., Meredith, E. K., Galichotte, E. N. et al. (2014). Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl. Acad. Sci. USA* **111**, 14888-14893.
- Choi, Y. and Mango, S. E. (2014). Hunting for Darwin's gemmules and Lamarck's fluid: transgenerational signaling and histone methylation. *Biochim. Biophys. Acta* **1839**, 1440-1453.

- Danchin, É., Charmantier, A., Champagne, F. A., Mesoudi, A., Pujol, B. and Blanchet, S. (2011). Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nat. Rev. Genet.* **12**, 475-486.
- Day, T. and Bonduriansky, R. (2011). A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. *Am. Nat.* **178**, E18-E36.
- Dias, B. G., Maddox, S. A., Klengel, T. and Ressler, K. J. (2015). Epigenetic mechanisms underlying learning and the inheritance of learned behaviors. *Trends Neurosci.* **38**, 96-107.
- Hunter, B., Hollister, J. D. and Bomblies, K. (2012). Epigenetic inheritance: what news for evolution? *Curr. Biol.* **22**, R54-R56.
- Irie, N., Weinberger, L., Tang, W. W. C., Kobayashi, T., Viukov, S., Manor, Y. S., Dietmann, S., Hanna, J. H. and Surani, M. A. (2015). SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253-268.
- Jablonka, E. and Lamb, M. J. (2015). The inheritance of acquired epigenetic variations. *Int. J. Epidemiol.*
- Kim, Y.-K., Yeo, J., Kim, B., Ha, M. and Kim, V. N. (2012). Short structured RNAs with low GC content are selectively lost during extraction from a small number of cells. *Mol. Cell* **46**, 893-895.
- Kozomara, A. and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68-D73.
- Liao, J.-Y., Ma, L.-M., Guo, Y.-H., Zhang, Y.-C., Zhou, H., Shao, P., Chen, Y.-Q. and Qu, L.-H. (2010). Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3' trailers. *PLoS ONE* **5**, e10563.
- Livnat, A. (2013). Interaction-based evolution: how natural selection and nonrandom mutation work together. *Biol. Direct* **8**, 24.
- Magnúsdóttir, E., Dietmann, S., Murakami, K., Günesdogan, U., Tang, F., Bao, S., Diamanti, E., Lao, K., Gottgens, B. and Azim Surani, M. (2013). A tripartite transcription factor network regulates primordial germ cell specification in mice. *Nat. Cell Biol.* **15**, 905-915.
- Mattick, J. S. (2012). Rocking the foundations of molecular genetics. *Proc. Natl. Acad. Sci. USA* **109**, 16400-16401.
- Melo, S. A., Sugimoto, H., O'Connell, J. T., Kato, N., Villanueva, A., Vidal, A., Qiu, L., Vitkin, E., Perelman, L. T., Melo, C. A. et al. (2014). Cancer exosomes perform cell-independent microRNA biogenesis and promote tumorigenesis. *Cancer Cell* **26**, 707-721.
- Morris, K. V. (2015). The theory of RNA-mediated gene evolution. *Epigenetics* **10**, 1-5.
- Noble, D. (2013). Physiology is rocking the foundations of evolutionary biology. *Exp. Physiol.* **98**, 1235-1243.
- Noble, D. (2015). Evolution beyond neo-Darwinism: a new conceptual framework. *J. Exp. Biol.* **218**, 7-13.
- Noble, D., Jablonka, E., Joyner, M. J., Müller, G. B. and Omholt, S. W. (2014). Evolution evolves: physiology returns to centre stage. *J. Physiol.* **592**, 2237-2244.
- Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* **465**, 721-727.
- Richards, E. J. (2008). Population epigenetics. *Curr. Opin. Genet. Dev.* **18**, 221-226.
- Rivoire, O. and Leibler, S. (2014). A model for the generation and transmission of variations in evolution. *Proc. Natl. Acad. Sci. USA* **111**, E1940-E1949.
- Santonocito, M., Vento, M., Guglielmino, M. R., Battaglia, R., Wahlgren, J., Ragusa, M., Barbegalio, D., Borzi, P., Rizzari, S., Maugeri, M. et al. (2014). Molecular characterization of exosomes and their microRNA cargo in human follicular fluid: bioinformatic analysis reveals that exosomal microRNAs control pathways involved in follicular maturation. *Fertil. Steril.* **102**, 1751-1761.e1.
- Sharma, A. (2014). Bioinformatic analysis revealing association of exosomal mRNAs and proteins in epigenetic inheritance. *J. Theor. Biol.* **357**, 143-149.
- Sharma, A. (2015a). Transgenerational epigenetic inheritance requires a much deeper analysis. *Trends Mol. Med.* **21**, 269-270.
- Sharma, A. (2015b). Transgenerational epigenetic inheritance: resolving uncertainty and evolving biology. *Biomol. Concepts* **6**, 87-103.
- Sharma, A. (2015c). Transgenerational epigenetic inheritance: emerging concepts and future prospects. *J. Reprod. Health Med.* **1**, 60-63.
- Skinner, M. K. (2015). Environmental epigenetics and a unified theory of the molecular aspects of evolution: a Neo-Lamarckian concept that facilitates Neo-Darwinian evolution. *Genome Biol. Evol.* **7**, 1296-1302.
- Smith, Z. D., Chan, M. M., Humm, K. C., Karnik, R., Mekhoubad, S., Regev, A., Eggen, K. and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611-615.
- Szyf, M. (2015). Nongenetic inheritance and transgenerational epigenetics. *Trends Mol. Med.* **21**, 134-144.
- Tang, W. W. C., Dietmann, S., Irie, N., Leitch, H. G., Floros, V. I., Bradshaw, C. R., Hackett, J. A., Chinnery, P. F. and Surani, M. A. (2015). A unique gene regulatory network resets the human germline epigenome for development. *Cell* **161**, 1453-1467.
- Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Pääbo, S. and Rebhan, M. and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457-466.
- Weick, E.-M. and Miska, E. A. (2014). piRNAs: from biogenesis to function. *Development* **141**, 3458-3471.
- Zhang, Y., Fan, M., Zhang, X., Huang, F., Wu, K., Zhang, J., Liu, J., Huang, Z., Luo, H., Tao, L. et al. (2014). Cellular microRNAs up-regulate transcription via interaction with promoter TATA-box motifs. *RNA* **20**, 1878-1889.

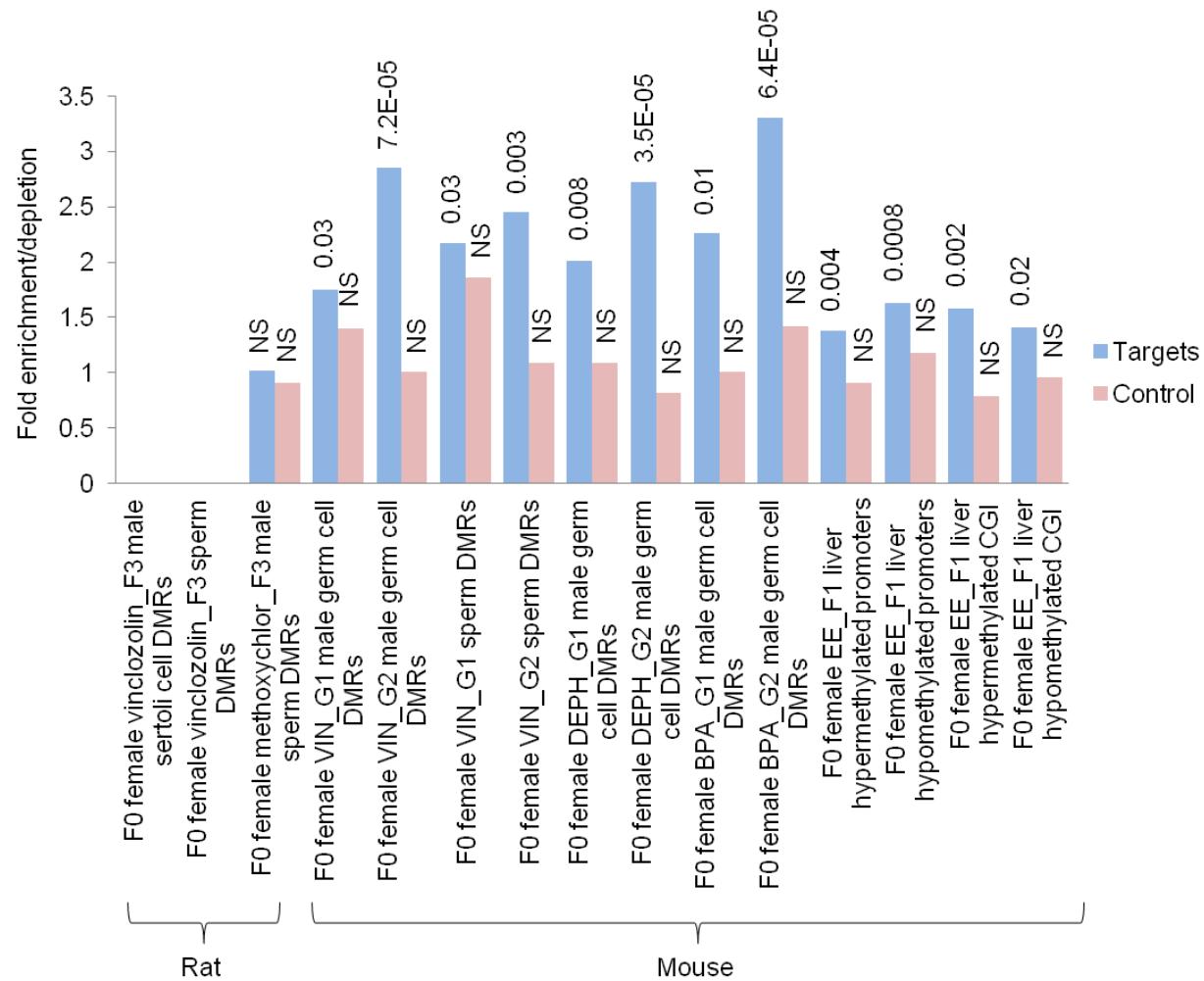


**Fig. S1. The concept of evolutionary transgenerational systems biology.**

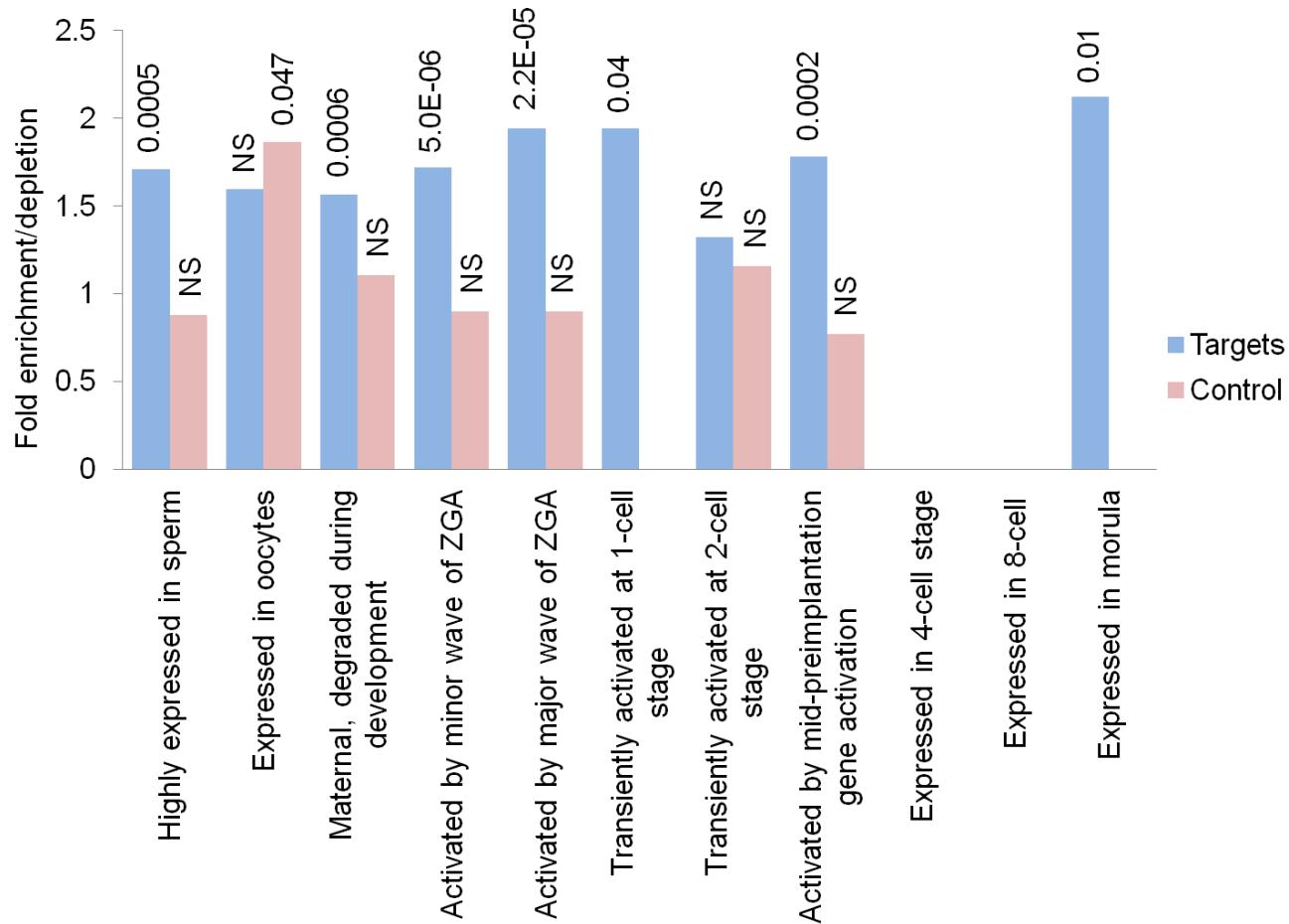
A representation of the conceptual mechanistic framework proposed previously (Sharma, 2014; Sharma, 2015a-c) to explain transgenerational epigenetic inheritance and its evolutionary significance. The model, supported by qualitative evidence, suggests that epigenetic inheritance of acquired traits may result in mutational fixation of the epiallele in evolutionary time course.



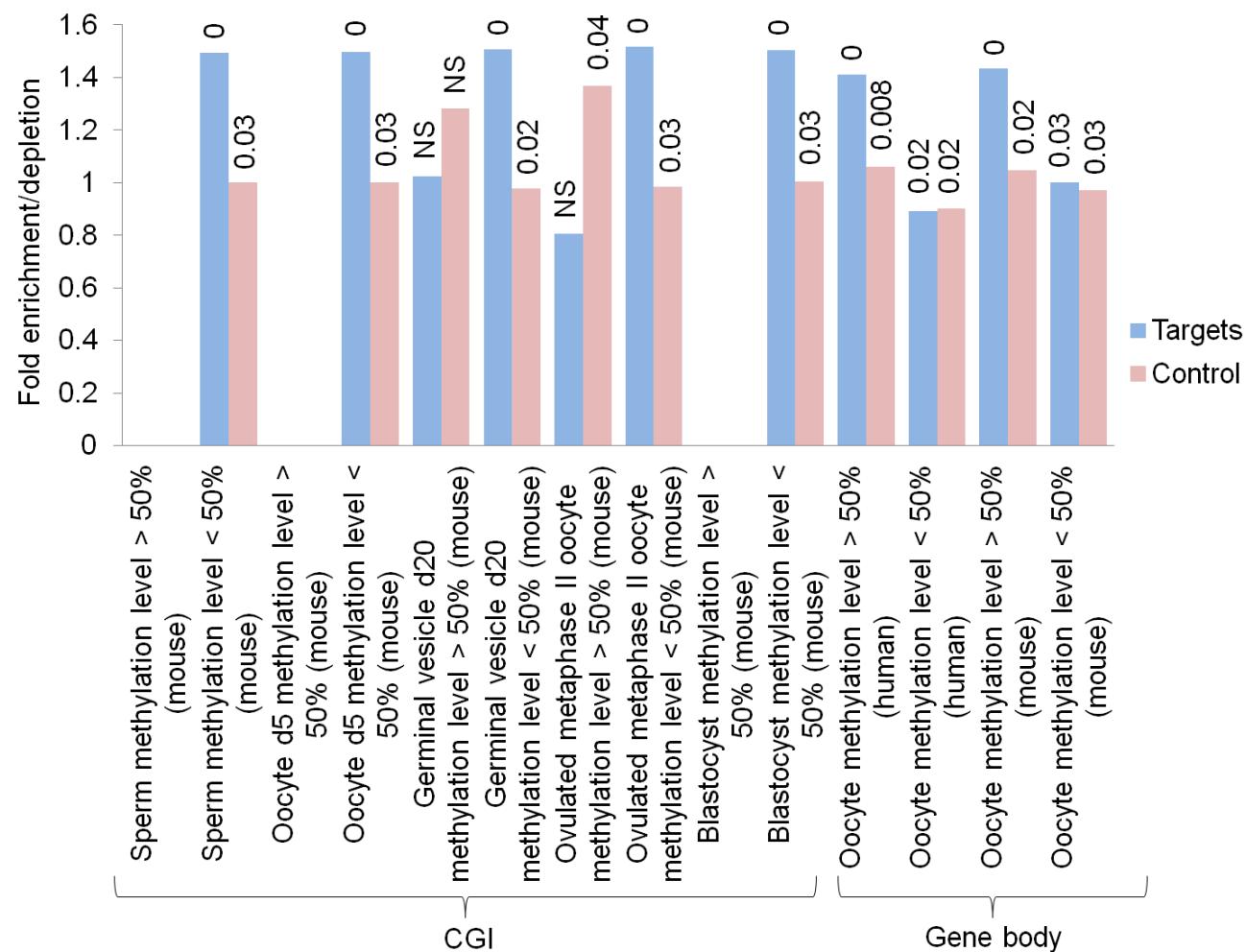
**Fig. S2. Enrichment of validated target genes of common miRNAs, total 708 in number (Table S8), in differentially expressed mRNAs reported in studies on epigenetic inheritance in rat.** Note combined analysis, represented by the extreme right pair of bars, showing enrichment of the targets in mRNAs reported in epigenetic studies. Key to mRNA expression profiles: gender exposed at F0-treatment at F0-generation, stage, gender and sample profiled. ME, male exposure; FE, female exposure; HFD, high fat diet; VIN, vinclozolin; PRD, protein restricted diet; HCD, high calorie diet. Enrichment *P* values, hypergeometric distribution, are shown above bars. NS, not significant. Missing bars indicate overlapping counts as either nil or excessively low, considered inappropriate for analysis.



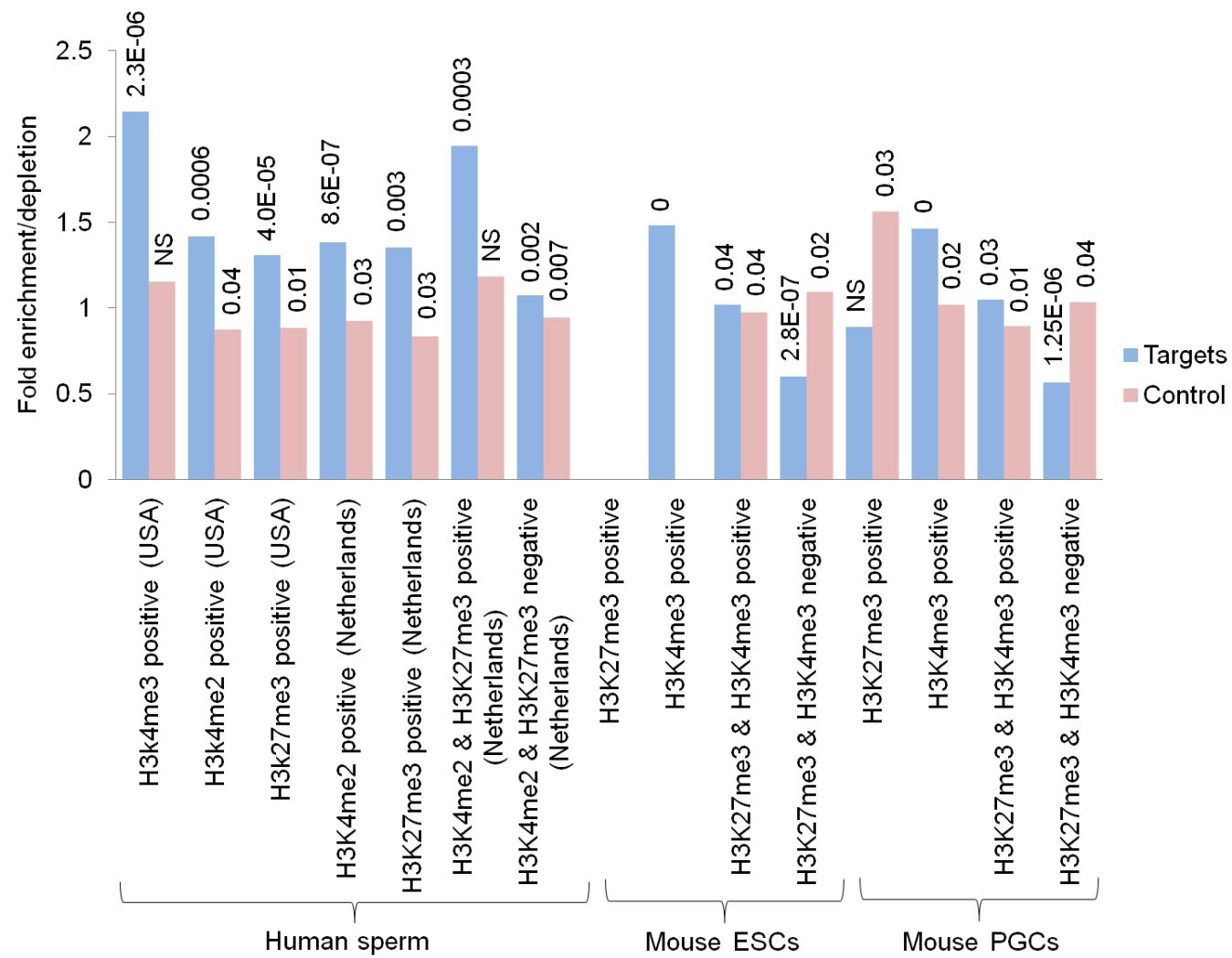
**Fig. S3.** Overrepresentation of targets (Fig. S2) in differentially methylated genes reported in studies on epigenetic inheritance. Key to DNA methylation profiles: gender and environmental factor used for exposure at F0-generation, gender, sample and DNA elements profiled. DEPH, di-(2-ethylhexyl)phthalate; EE, environmental enrichment; DMRs, differentially methylated regions; CGI, CpG island. Enrichment *P* values, hypergeometric distribution, are shown above bars. NS, not significant. Missing bars indicate overlapping counts as either nil or excessively low, considered inappropriate for analysis.



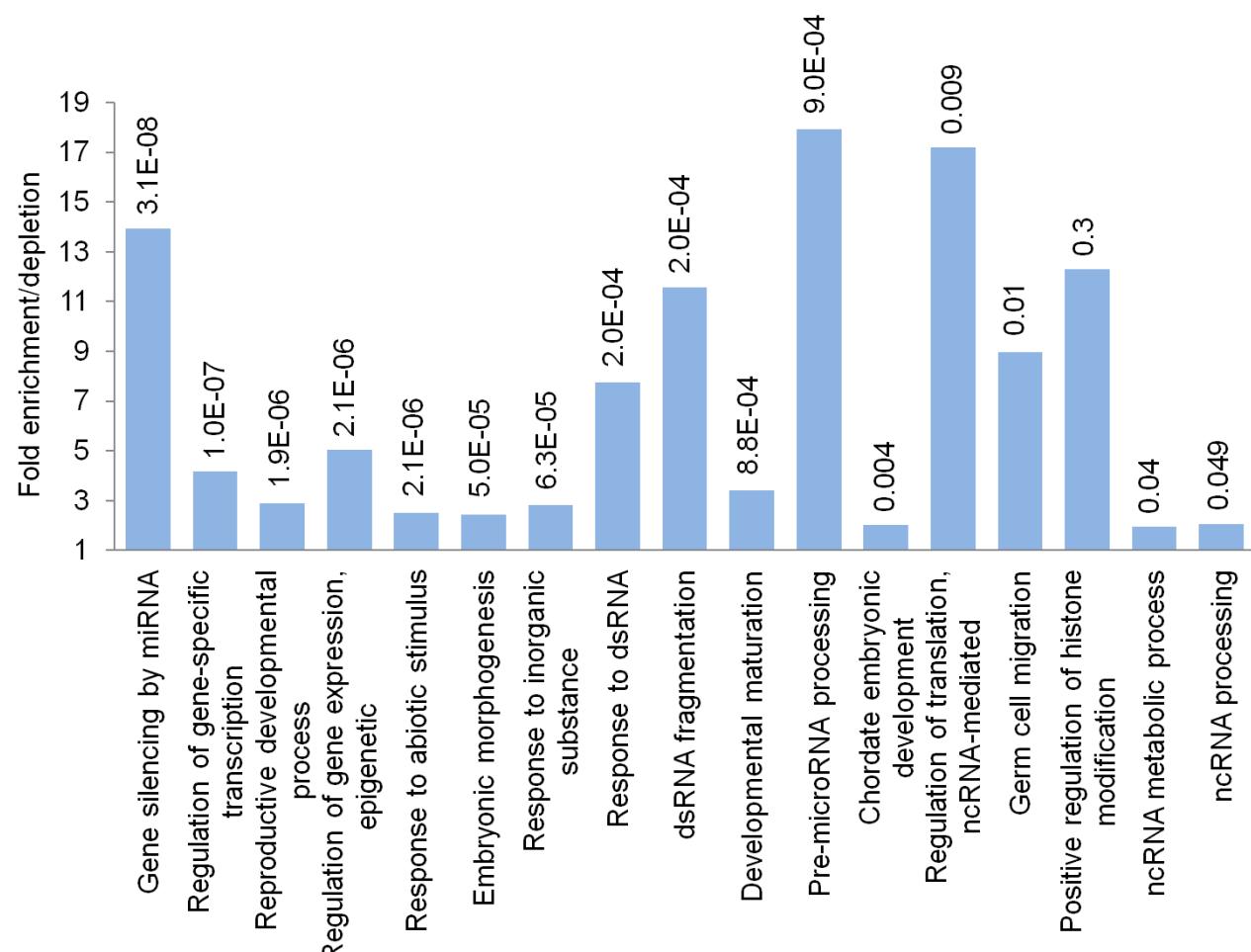
**Fig. S4. Gene expression in gametes, zygote and postzygotic stages.** Overall, the targets tend to enrich genes expressed in sperm, oocytes, zygote and developing embryo. ZGA, zygotic gene activation. Enrichment *P* values, hypergeometric distribution, are shown above bars. NS, not significant. Missing bars indicate overlapping counts as either nil or excessively low, considered inappropriate for analysis.



**Fig. S5. DNA methylation levels in gametes and developing embryo.** The targets in general overrepresent genes with < 50%, and underepresent genes with > 50% CpG island (CGI) methylation levels. For gene body levels, the targets enrich genes with > 50% methylation, without showing any trend for over- or under-representation of genes with < 50% levels. Enrichment *P* values, hypergeometric distribution, are shown above bars. Values with 16 or more negative exponent of 10 were rounded down to zero. NS, not significant. Missing bars indicate overlapping counts as either nil or excessively low, considered inappropriate for analysis.



**Fig. S6. Histone modifications in gametes and development.** Note clear enrichment of the targets in H3K4me2, H3K4me3, and H3K27me3 positive genes in human sperm. The targets also show differential histone modifications in embryonic stem cells (ESCs) and PGCs. Enrichment *P* values, hypergeometric distribution, are shown above bars. Values with 16 or more negative exponent of 10 were rounded down to zero. NS, not significant. Missing bars indicate overlapping counts as either nil or excessively low, considered inappropriate for analysis.



**Fig. S7. Gene ontology biological process enrichment.** A subset of processes enriched in the targets is shown. Control genes did not show enrichment/depletion of any process, as expected. Enrichment *P* values, hypergeometric distribution, are shown above bars.

**Table S1**

[Click here to Download Table S1](#)

**Table S2**

[Click here to Download Table S2](#)

**Table S3**

[Click here to Download Table S3](#)

**Table S4**

[Click here to Download Table S4](#)

**Table S5.** Data source used in the analysis shown in Fig. 2 and Figs. S2-S7

<i>Figure label</i>	<i>Data source</i>
ME-HFD-F0 testis	PubMed ID: 23845863
ME-LPD-F1 liver	PubMed ID: 21183072
ME-B[a]P-F1 embryo 2 cell	PubMed ID: 22548065
ME-B[a]P-F1 embryo 8 cell	PubMed ID: 22548065
ME-B[a]P-F1 embryo blastocyst	PubMed ID: 22548065
FE-stress-F1 sperm	PubMed ID: 24728267
ME-stress-F1 brain	PubMed ID: 23699511
ME-stress-F0 sperm	PubMed ID: 23699511
FE-BA-F1 fetal d 65 ovary	PubMed ID: 23525218
FE-BA-F1 fetal d 90 ovary	PubMed ID: 23525218
FE-T-F1 fetal d 90 ovary	PubMed ID: 22009729
FE-TF-F1 fetal d 90 ovary	PubMed ID: 22009729
FE-T vs. TF-F1 fetal d 90 ovary	PubMed ID: 22009729
FE-HCD-F0 adult female uterus	PubMed ID: 21862610
ME-HFD-F1 adult female islets	PubMed ID: 20962845
FE-VIN-F1 embryo male testis	PubMed ID: 18042343
FE-VIN-F1 embryo E13 male testis	PubMed ID: 20566332
FE-VIN-F1 embryo E14 male testis	PubMed ID: 20566332
FE-VIN-F1 embryo E16 male testis	PubMed ID: 20566332
FE-PRD-F1 adult female liver	PubMed ID: 21750721
FE-HCD-F1 embryo male blastocyst	PubMed ID: 21862610
FE-VIN-F2 embryo male testis	PubMed ID: 18042343
FE-PRD-F2 adult female liver	PubMed ID: 21750721
FE-VIN-F3 embryo male testis	PubMed ID: 18042343
FE-PRD-F3 adult female liver	PubMed ID: 21750721
FE-VIN-F3 adult male sertoli cells	PubMed ID: 23555832
FE-VIN-F3 adult female granulosa cells	PubMed ID: 22570695
FE-VIN-F3 adult female heart	PubMed ID: 23034163
FE-VIN-F3 adult female kidney	PubMed ID: 23034163
FE-VIN-F3 adult female liver	PubMed ID: 23034163
FE-VIN-F3 adult female uterus	PubMed ID: 23034163
FE-VIN-F3 adult male heart	PubMed ID: 23034163
FE-VIN-F3 adult male kidney	PubMed ID: 23034163
FE-VIN-F3 adult male liver	PubMed ID: 23034163
FE-VIN-F3 adult male prostate	PubMed ID: 23034163
FE-VIN-F3 adult male seminal vesicle	PubMed ID: 23034163
FE-VIN-F3 adult female ovary	PubMed ID: 23034163

FE-VIN-F3 adult male testis	PubMed ID: 23034163
FE-VIN-F3 adult male amygdala	PubMed ID: 19015723
FE-VIN-F3 adult male hippocampus	PubMed ID: 19015723
FE-VIN-F3 adult female amygdala	PubMed ID: 19015723
FE-VIN-F3 adult female hippocampus	PubMed ID: 19015723
FE-VIN-F3 adult male brain	PubMed ID: 19015723
FE-VIN-F3 adult male basolateral amygdala	PubMed ID: 22615374
FE-VIN-F3 adult male brain cortex	PubMed ID: 22615374
FE-VIN-F3 adult male hippocampus CA1	PubMed ID: 22615374
FE-VIN-F3 adult male hippocampus CA3	PubMed ID: 22615374
FE-VIN-F3 adult female basolateral amygdala	PubMed ID: 22615374
FE-VIN-F3 adult female brain cortex	PubMed ID: 22615374
FE-VIN-F3 adult female hippocampus CA1	PubMed ID: 22615374
FE-VIN-F3 adult female hippocampus CA3	PubMed ID: 22615374
FE-VIN-F3 adult male ventral prostate	PubMed ID: 18220299
FE-VIN-F3 adult male prostate epithelial cells	PubMed ID: 18220299
FE-VIN-F3 embryo E13 male primordial germ cells	PubMed ID: 23869203
FE-VIN-F3 embryo E16 male prospermatogonia	PubMed ID: 23869203
F0 female vinclozolin_F3 male sertoli cell DMRs	PubMed ID: 23555832
F0 female vinclozolin_F3 sperm DMRs	PubMed ID: 20927350
F0 female methoxychlor_F3 male sperm DMRs	PubMed ID: 25057798
F0 female VIN_G1 male germ cell DMRs	PubMed ID: 25853433
F0 female VIN_G2 male germ cell DMRs	PubMed ID: 25853433
F0 female VIN_G1 sperm DMRs	PubMed ID: 25853433
F0 female VIN_G2 sperm DMRs	PubMed ID: 25853433
F0 female DEPH_G1 male germ cell DMRs	PubMed ID: 25853433
F0 female DEPH_G2 male germ cell DMRs	PubMed ID: 25853433
F0 female BPA_G1 male germ cell DMRs	PubMed ID: 25853433
F0 female BPA_G2 male germ cell DMRs	PubMed ID: 25853433
F0 female EE_F1 liver hypermethylated promoters	PubMed ID: 25853433
F0 female EE_F1 liver hypomethylated promoters	PubMed ID: 25853433
F0 female EE_F1 liver hypermethylated CGI	PubMed ID: 25853433
F0 female EE_F1 liver hypomethylated CGI	PubMed ID: 25853433
Imprinted (Human)	<a href="http://www.geneimprint.com/">http://www.geneimprint.com/</a>
Predicted (Human)	<a href="http://www.geneimprint.com/">http://www.geneimprint.com/</a>
Not imprinted (Human)	<a href="http://www.geneimprint.com/">http://www.geneimprint.com/</a>
Imprinted (Mouse)	<a href="http://www.geneimprint.com/">http://www.geneimprint.com/</a>
Predicted (Mouse)	<a href="http://www.geneimprint.com/">http://www.geneimprint.com/</a>
Not imprinted (Mouse)	<a href="http://www.geneimprint.com/">http://www.geneimprint.com/</a>
Imprinted (Rat)	<a href="http://www.geneimprint.com/">http://www.geneimprint.com/</a>
mRNA (Human)	<a href="http://ctdbase.org/">http://ctdbase.org/</a>

protein (Human)	<a href="http://ctdbase.org/">http://ctdbase.org/</a>
mRNA (Mouse)	<a href="http://ctdbase.org/">http://ctdbase.org/</a>
protein (Mouse)	<a href="http://ctdbase.org/">http://ctdbase.org/</a>
mRNA (Rat)	<a href="http://ctdbase.org/">http://ctdbase.org/</a>
protein (Rat)	<a href="http://ctdbase.org/">http://ctdbase.org/</a>
Highly expressed in sperm	PubMed ID: 23471003
Expressed in oocytes	PubMed ID: 23892778
Maternal, degraded during development	<a href="http://dbtmee.hgc.jp/">http://dbtmee.hgc.jp/</a>
Activated by minor wave of ZGA	<a href="http://dbtmee.hgc.jp/">http://dbtmee.hgc.jp/</a>
Activated by major wave of ZGA	<a href="http://dbtmee.hgc.jp/">http://dbtmee.hgc.jp/</a>
Transiently activated at 1-cell stage	<a href="http://dbtmee.hgc.jp/">http://dbtmee.hgc.jp/</a>
Transiently activated at 2-cell stage	<a href="http://dbtmee.hgc.jp/">http://dbtmee.hgc.jp/</a>
Activated by mid-preimplantation gene activation	<a href="http://dbtmee.hgc.jp/">http://dbtmee.hgc.jp/</a>
Expressed in 4-cell stage	PubMed ID: 23892778
Expressed in 8-cell	PubMed ID: 23892778
Expressed in morula	PubMed ID: 23892778
Sperm methylation level > 50% (mouse)	PubMed ID: 21706000
Sperm methylation level < 50% (mouse)	PubMed ID: 21706000
Oocyte d5 methylation level > 50% (mouse)	PubMed ID: 21706000
Oocyte d5 methylation level < 50% (mouse)	PubMed ID: 21706000
Germinal vesicle d20 methylation level > 50% (mouse)	PubMed ID: 21706000
Germinal vesicle d20 methylation level < 50% (mouse)	PubMed ID: 21706000
Ovulated metaphase II oocyte methylation level > 50% (mouse)	PubMed ID: 21706000
Ovulated metaphase II oocyte methylation level < 50% (mouse)	PubMed ID: 21706000
Blastocyst methylation level > 50% (mouse)	PubMed ID: 21706000
Blastocyst methylation level < 50% (mouse)	PubMed ID: 21706000
Oocyte methylation level > 50% (human)	PubMed ID: 25501653
Oocyte methylation level < 50% (human)	PubMed ID: 25501653
Oocyte methylation level > 50% (mouse)	PubMed ID: 25501653
Oocyte methylation level < 50% (mouse)	PubMed ID: 25501653
H3k4me3 positive (USA)	PubMed ID: 19525931
H3k4me2 positive (USA)	PubMed ID: 19525931
H3k27me3 positive (USA)	PubMed ID: 19525931
H3K4me2 positive (Netherlands)	PubMed ID: 20473313
H3K27me3 positive (Netherlands)	PubMed ID: 20473313
H3K4me2 & H3K27me3 positive (Netherlands)	PubMed ID: 20473313
H3K4me2 & H3K27me3 negative (Netherlands)	PubMed ID: 20473313
H3K27me3 positive	PubMed ID: 23727241
H3K4me3 positive	PubMed ID: 23727241

H3K27me3 & H3K4me3 positive	PubMed ID: 23727241
H3K27me3 & H3K4me3 negative	PubMed ID: 23727241
H3K27me3 positive	PubMed ID: 23727241
H3K4me3 positive	PubMed ID: 23727241
H3K27me3 & H3K4me3 positive	PubMed ID: 23727241
H3K27me3 & H3K4me3 negative	PubMed ID: 23727241
Expressed in all tissues	PubMed ID: 25613900
Tissue enriched	PubMed ID: 25613900
Group enriched	PubMed ID: 25613900
Tissue enhanced	PubMed ID: 25613900
Total proteins	PubMed ID: 25613900
Low CpG promoters	PubMed ID: 17334365
Intermediate CpG promoters	PubMed ID: 17334365
High CpG promoters	PubMed ID: 17334365
All cell/tissue types	PubMed ID: 25693566
Cell/tissue types excluding ESCs	PubMed ID: 25693566

Table S6

[Click here to Download Table S6](#)

**Table S7. Data source used in the analysis shown in Fig. 3.**

<i>Figure label</i>	<i>Data source</i>
LCP	PubMed ID: 22016335
HCP	PubMed ID: 22016335
NHGRI	<a href="https://www.genome.gov/">https://www.genome.gov/</a>
GAD	<a href="http://geneticassociationdb.nih.gov/">http://geneticassociationdb.nih.gov/</a>
MIM	<a href="http://www.omim.org/">http://www.omim.org/</a>
LSDB	<a href="http://geneticassociationdb.nih.gov/">http://geneticassociationdb.nih.gov/</a>
CTD	<a href="http://ctdbase.org/">http://ctdbase.org/</a>

Table S8

[Click here to Download Table S8](#)