Multi-omics data analysis supporting inheritance of acquired traits in humans

Abhay Sharma

CSIR-Institute of Genomics and Integrative Biology, Sukhdev Vihar, Mathura Road, New Delhi 110025, India. Tel: +91-11-26932421. Email: abhaysharma@igib.res.in

ORCID 0000-0003-1306-1497

Abstract

Emerging evidence increasingly supports germline mediated non-DNA sequence-based inheritance of environmentally induced traits in animals including mammalian models. Epidemiological and ancestral exposure studies also suggest possible presence of epigenetic inheritance in humans. Notably, cross-species conservation of phenotypic responses in non-genetic inheritance has recently been shown in transcriptomic analysis of *Drosophila* and rodent models. Here, an unbiased analysis of available human and animal model multi-omics data is presented that supports inheritance of induced traits in humans. First, in an analysis feasibility survey, physical exercise was identified as the lone environmental factor that is known to affect conserved physiological pathways and to induce inheritance in animal models, and at the same time, the effects of which on human cohort sperm DNA methylome, and human cohort, rat, mouse and *Drosophila* coding and/or non-coding somatic transcriptome have also been multiply described. The hypothesis was thus tested that exercise induced human sperm DNA methylation changes can specifically explain exercise induced human and animal model somatic transcriptome changes. For controlling the analysis, bariatric surgery, identified as the lone environmental factor for which human cohort sperm DNA methylome and somatic transcriptome datasets are multiply available, was used. Remarkably, the hypothesis was supported in all pre-specified tests, involving gene set overlap, gene ontology enrichment, promotor methylation and gene regulation directionalities, and coding and non-coding transcriptome interactions. The results presented support germline encoding of acquired traits in humans, thus suggesting existence of epigenetic inheritance in our species.

Introduction

Accumulating evidence suggests that information about induced phenotypic traits in animals including mammals can be non-genetically transmitted across organismal generations through the germline (Cavalli and Heard, 2019; Perez and Lehner, 2019; Anastasiadi *et al*., 2021; Fitz-James and Cavalli, 2022; Santilli and Boskovic, 2023). This unconventional, non-DNA sequence-based inheritance is commonly referred to as transgenerational epigenetic inheritance (TEI), notwithstanding a general lack of knowledge about the exact germline factor(s) inherited. Germline involvement is assumed mainly because TEI calling requires phenotypic appearance in descendants at least one generation away from that which could have possibly been affected by direct exposure of inheritance triggering elements in the founders. In mammals, for example, the requirement is observation of a phenotype at least two generations away from the founders when the latter are males, and three generations when gestating females (Cavalli and Heard, 2019; Perez and Lehner, 2019; Fitz-James and Cavalli, 2022). Similarly, epigenetic involvement is assumed because genetic variations causing inheritance of traits that are triggered by environmental agents seems unlikely, especially given use of agents often not known to cause mutations, and of inbred and isogenic strains that minimize genetic background differences. Instead, epigenetic components like DNA methylation, histone post-translational modifications, and RNA are mainly considered as potential carriers or mediators of phenotypic information across generations (Cavalli and Heard, 2019; Perez and Lehner, 2019; Anastasiadi *et al*., 2021; Santilli and Boskovic, 2023; Takahashi *et al*., 2023). Besides, intergenerational inheritance, that does not exclude the possibility of direct exposure effects but nonetheless shares features with the former, is considered relevant in epigenetic inheritance. Seemingly contradicting multiple fundamental concepts including neo-Darwinian principle that limits the source of heredity exclusively to DNA sequence, Weismann barrier that prohibits soma to germline transfer of hereditary information, and developmental reprogramming that inhibits

germline inheritance of epigenetic signals, inheritance of acquired traits has intriguing implications in ecology and evolution, and health and disease (Cavalli and Heard, 2019; Anastasiadi *et al*., 2021; Fitz-James and Cavalli, 2022; Santilli and Boskovic, 2023; Adrian-Kalchhauser et al., 2020). Environmentally induced TEI, first described as a consequence of exposure of gestating female rats to an endocrine disrupting compound (Anway *et al*., 2005), has been shown in invertebrates and vertebrates alike, with triggers including toxicants and pollutants, dietary factors, and stress, besides others (Bohacek and Mansuy, 2017; Cavalli and Heard, 2019; Perez and Lehner, 2019; Fitz-James and Cavalli, 2022; Santilli and Boskovic, 2023). However, difficulties in eliminating confounders like genetic differences especially in rat models lacking inbred stocks, maternal contribution, and *in utero* and postnatal effects in female exposure models remain a concern (Bohacek and Mansuy, 2017; Cavalli and Heard, 2019). Given this, paternal exposure and patrilineal models of inheritance involving founders with similar genetic background are considered relatively less confounding and hence the focus of recent investigations (Bohacek and Mansuy, 2017; Cavalli and Heard, 2019; Perez and Lehner, 2019). With environmentally induced transgenerational spermatogenic inheritance first reported in the fruit fly *Drosophila* using a neuroactive compound as the trigger in an isogenic fly line (Sharma and Singh, 2009), male exposure and male line mediated inheritance, inter- or trans-generational, has been shown in various species including mammals (Bohacek and Mansuy, 2017; Perez and Lehner, 2019; Fitz-James and Cavalli, 2022; Santilli and Boskovic, 2023).

Epidemiological, and ancestral exposure studies also suggest potential presence of epigenetic inheritance in humans (Senaldi and Smith-Raska, 2020; Vågerö *et al*., 2022; Kaufman et al., 2023; Niu *et al*., 2023; Santilli and Boskovic, 2023). Similar to the findings in diet induced paternal intergenerational inheritance in animal models, dietary intervention has been associated with sperm non-coding RNA content change in humans, thus supporting possible

existence of germline mediated non-genetic inheritance in the latter (Nätt *et al*., 2019). However, considering a lack of robust molecular evidence linking induced traits and germline responses, the concern has been raised that assuming existence of epigenetic inheritance in humans due to supporting findings in rodents runs the risk of being unethical and societally harmful, as, for example, it can unnecessarily cause anxiety in prospective parents by imposing unrealistic demands (Juengst *et al*., 2014). The obvious difficulties (Senaldi and Smith-Raska, 2020; Santilli and Boskovic, 2023) in conducting formal human non-genetic germline inheritance studies aside, integrative analysis of diverse data types may seem to provide a valuable general strategy for understanding existential, mechanistic, evolutionary, and biomedical aspects of epigenetic inheritance (Molaro *et al*., 2011; Cavalli and Heard, 2019; Sharma, 2020; Niu *et al*., 2023). Further dissecting the already described animal models of non-genetic inheritance, and leveraging stronger evidence of that in invertebrates to assess mammalian model evidence are examples of suggested approaches (Perez and Lehner, 2019; Fitz-James and Cavalli, 2022; Bhalla and Sharma, 2022). Notably, despite expected differences in reproductive and epigenetic mechanisms among diverse species, paternal non-genetic inheritance induced by similar metabolic triggers in *Drosophila* and murine models has recently been associated with similar transcriptomic changes across species (Bhalla and Sharma, 2023). This catalyses the need for an unbiased integrative analysis of available omics scale relevant human and animal model data towards closely examining whether evidence supports inheritance of environmentally induced traits in humans. An analysis is presented here to that end. First, a feasibility survey was conducted to assess if required data is available to perform the concept analysis. On finding the availability, an analysis was designed, and based on the design, the data collected and processed, and the results of the pre-specified analysis obtained. The findings clearly support inheritance of induced traits in humans.

Feasibility survey

The concept analysis assumed availability of studies reporting, separately or in combination, environmental factor matched omics level effects in animal models and human cohorts using a self-controlled, pre- and post-exposure design, provided the factor(s) is known to induce male line mediated inter- or trans-generational inheritance in paternal exposure animal models. Studies employing a pre-post study design were chosen in order to minimize inter-individual variability in environmental effects (Wang *et al*., 2005; Gallego-Paüls *et al*., 2021). Availability of the target studies was explored by searching PubMed and Gene Expression Omnibus (GEO) databases of the National Center for Biotechnology Information (Sayers *et al*., 2022). The search led to the finding that physical exercise was a factor that is known to induce paternal intergenerational inheritance in mice, besides maternal in both mice and rats (Kusuyama *et al*., 2020; Vieira de Sousa Neto *et al*., 2021), and for which diverse sets of human sperm DNA methylome, and coding and non-coding transcriptome as well as rat, mouse, and *Drosophila* coding transcriptome data are available. Besides established rat and mouse models, fly has also recently emerged as a model of exercise biology relevant in mammals and highly appropriate for understanding genetic architecture of exercise associated traits (Watanabe and Riddle, 2019). All these three exercise models were thus considered. Together, these positive search results assured that a basic analysis can be designed wherein exercise associated human sperm DNA methylome can be compared, for gene set similarities, with exercise associated human somatic transcriptome, and to validate the findings, with exercise associated animal model somatic transcriptome. Next, to add value in analysis design, PubMed and GEO were again searched for finding human cohort studies that could together provide non-exercise treatment matched sperm DNA methylome and somatic tissue transcriptome data to serve as controls. This search identified bariatric surgery as a treatment for which both the data types existed. Both exercise (Ingerslev *et al*., 2018) and bariatric surgery (Donkin *et al*., 2016) human sperm

methylome studies primarily focused on the respective experimental findings, providing limited integrative insights. Cumulatively, the feasibility survey confirmed the availability of data required for performing the concept analysis.

Analysis design

Considering the extent of data availability, an analysis was designed to first test whether or not exercise associated differentially methylated genes (DMGs) in human sperm specifically overrepresent exercise associated differentially expressed mRNAs (DEmRNAs) in human somatic tissues. The underlying hypothesis was that an intervention specific overrepresentation would support germline epigenetic encoding of gene expression in soma, and hence inheritance of induced traits. Considering that skeletal muscle is the main target tissue of exercise (Vina *et al*., 2012), the same was to be essentially included among somatic tissues. With these, supporting evidence was to lead to testing if exercise associated human sperm DMGs also overrepresent exercise associated DEmRNAs in human somatic tissue matched animal model samples. Favourable findings were to be followed by examining if human DMGs and DEmRNAs, and animal model DEmRNAs show common enrichment for gene ontology biological process terms relevant in exercise biology. Affirmatory results were to lead to focused human data analysis, testing first whether or not the directionality of promoter DMGs in sperm shows a trend that is opposite of that of somatic DEmRNAs, because promoter methylation is in general considered to inversely regulate gene expression (Pepin *et al*., 2019). Next, on finding confirmatory results, given that miRNAs are considered as essential mediators of exercise associated processes (Silva *et al*., 2017), it was to be examined if the sperm promoter DMGs are enriched for targets of reported exercise associated differentially expressed miRNAs (DEmiRNAs) in soma in genome scale studies. Finally, supporting results were to lead to testing whether or not those miRNA target genes more frequently occur as

DEmRNA genes, excluding the ones present among sperm DMGs, in somatic tissues than expected by chance.

Data collection

The DMGs, DEmRNAs, and DEmiRNAs were identified using <0.05 adjusted *p* value as significance level. For DMGs, full sets of original author-identified genes annotating to differentially methylated regions were considered, irrespective of percentage of methylation difference. DMGs showing both hyper- and hypo-methylation at different chromosomal regions were not considered. For DEmRNAs and DEmiRNAs, >1.2 fold change was required for inclusion. Complete sets of DEmRNAs were identified from original reports directly, wherever provided, or from the corresponding GEO submissions, mainly using GEO2R (Barrett *et al*., 2013) and GREIN (Mahi *et al*., 2019), for microarray and RNA-seq data, in that order. Duplicate DEmRNAs, if any, were removed in excel after alphabetically sorting the gene symbols. Original author-identified complete sets of DEmiRNAs were directly used. The gene symbols were harmonized using, as appropriate, HumanMine (Lyne *et al*., 2022), rat and mouse options in MouseMine (Motenko *et al*., 2015), and FlyMine (Lyne *et al*., 2007) first, and then to HGNC using Multi-symbol checker (Seal et al., 2023), RGD and MGI using g:Convert (Reimand *et al*., 2007), and FlyBase using ID Validator (Gramates *et al*., 2022). The miRNA nomenclature in original studies was directly used. A compendium of all DMGs, DEmRNAs, and DEmiRNAs finally used in the analysis, along with sample information, publication and data source references, as also labels assigned here to individual samples for identifying them in the result section are provided in **Supplementary Table 1**.

Data processing

Gene set overlap was examined using hypergeometric distribution with sampling without replacement from a finite population, considering a population size of 25000 human genes. The

rat and mouse gene symbols in upper case were used for finding enrichment between rodent and human gene sets. For comparing *Drosophila* and human gene sets, human orthologs of fly genes were identified using DIOPT Ortholog Finder (Hu *et al*., 2011), irrespective of rank and scores. Enrichment of gene ontology biological process terms was determined using Gene Ontology Resource (Gene Ontology Consortium, 2021), with all human genes as the default reference, Fisher's exact test and <0.05 FDR as the statistical criteria, and >1.2 fold enrichment as the requirement. For gene ontology analysis, human, rat, and mouse gene symbols were used directly, while human orthologs of fly genes were considered. Enrichment of miRNA target genes was found using ToppGene Suite (Chen *et al*., 2009), with all human genes as default background, and <0.05 Bonferroni *q*-value as significance threshold. The miRNA targets were identified using TargetScan (Agarwal *et al*., 2015), irrespective of aggregate P_{CT} and cumulative weighted context score. The chi-square goodness of fit test was performed to compare observed and expected probability distribution. Random numbers from RANDOM.ORG (Haahr, 2023) were used for gene randomization.

The DMGs, DEmRNAs, DEmiRNAs, and gene ontology enrichment were visualized using, as appropriate, methylation difference, $\log_2$ fold change, and $\log_2$ fold enrichment based heatmaps generated in Heatmapper (Babicki *et al*., 2016), with the scale type none, and/or the clustering method average linkage and the distance measurement method Euclidean. A reduced number of heatmap shades was used to discretize hyper- and hypo-methylation in DMGs, and up- and down-regulation in DEmRNAs. Bar graphs were used to display $\log_2$ fold enrichment in gene set overlaps, and gene ontology enrichment in miRNA targets as well as to exhibit $-\log_{10} q$ value for miRNA target enrichment.

Results and Discussion

Results of the above processing of the collected data are presented and discussed below, in the order pre-specified in the analysis design.

Human exercise sperm DMGs specifically overrepresent human exercise somatic DEmRNAs

It was of the foremost interest to examine the overlap of exercise associated human sperm DMGs and of bariatric surgery associated human sperm DMGs with human somatic tissue DEmRNAs associated with the two interventions. Separate clustering of exercise associated sperm DMGs, and skeletal muscle and blood DEmRNAs, and bariatric surgery associated sperm DMGs and subcutaneous adipose tissue DEmRNAs showed largely similar regulation between samples and studies related to a given intervention (**Fig. 1a-e**). Given this directional consistency, all the genes belonging to different samples within a cluster were pooled for gene overlap analysis, as greater gene set size would increase statistical strength of overlap tests. To avoid known gene set size related biases in enrichment analysis (Wang *et al.*, 2010) and enable direct comparison between exercise sperm DMGs and bariatric surgery sperm DMGs, the genes in the latter set, that was almost 10 times higher in number compared to the genes in the former, was randomized and then divided into 10 subsets of similar sizes. A total of 11 sperm DMG lists, one related to exercise and the rest to bariatric surgery, were thus finally tested for overlap with DEmRNA gene lists. Remarkably, the exercise related sperm DMGs, compared to bariatric surgery related sperm DMGs, showed unequivocal higher enrichment of exercise, not bariatric surgery, related somatic DEmRNAs (**Fig. 1f**). On the other hand, the reverse, bariatric surgery related sperm DMGs showing unequivocal higher abundance of bariatric surgery related DEmRNAs, was not observed. These results specifically demonstrated that exercise associated sperm DMGs and somatic DEmRNAs are of similar gene regulatory

significance, thus supporting the plausibility of inheritance of exercise induced traits in humans.

Human exercise sperm DMGs specifically overrepresent animal model exercise DEmRNAs

The above human alone comparison was next repeated for comparing exercise and bariatric surgery associated human sperm DMGs with exercise associated DEmRNAs in animal models. The animal data included DEmRNAs in skeletal muscle in rats and mice, and in larval body wall muscles and whole organism adult fly homogenate in *Drosophila*. On previous lines, species specific clustering of DEmRNAs showed an overall trend for directional similarity in gene expression regulation across samples and studies (**Fig. 2a-c**). As above, all the DEmRNAs belonging to a given species were pooled for overlap tests. Strikingly, the exercise related human sperm DMGs, in comparison to that related to bariatric surgery, showed unambiguous higher enrichment of exercise related DEmRNAs in rat, mouse, and fly models alike (**Fig. 2d**). These results provided further support for exercise induced epigenetic inheritance in humans.

Human exercise sperm DMGs are enriched for exercise related gene ontology terms

Next, all exercise associated human and animal model gene sets used for gene overlap analysis were subjected to gene ontology analysis for enrichment of biological process terms. Overall, various terms were commonly enriched among these gene sets (**Fig. 3a**). Strikingly, muscle structure development showed highest enrichment in sperm DMGs, with the term also showing high enrichment in human, rat, and mouse skeletal muscle DEmRNAs, and larval body wall muscles and whole organism fly homogenate DEmRNAs combined for *Drosophila*. Human blood DEmRNAs alone did not show enrichment of this term. A survey of all the enriched, “muscle” containing terms in all the gene sets revealed enrichment of a multiplicity of those terms in sperm DMGs as well as in all except blood DEmRNA gene sets (**Fig. 3b**). These

results further supported the biological plausibility of exercise induced non-genetic inheritance in humans. Hereafter, the analysis focused on human data alone.

Human exercise sperm promoter DMGs and somatic DEmRNAs show inverse regulation

Whether exercise associated human sperm promoter DMGs and exercise associated human somatic tissue DEmRNAs are oppositely associated in terms of directionality was investigated next. The DMGs common to both the individual sperm samples were separated in two groups, promoter and non-promoter, with the latter group representing intron, exon, downstream, distal intergenic, and 5' UTR combined, and each group compared with all the individual exercise DEmRNA samples as test and all the individual bariatric surgery DEmRNA samples as control (**Fig. 4a**). Startlingly, chi-square goodness of fit test specifically showed that the observed counts of DEmRNA events showing opposite of DMG directionality in the exercise samples were significantly higher than the expected counts assuming equal chance for DEmRNA events to exhibit same and opposite regulation (**Fig. 4b**). These findings provided crucial, additional evidence supporting exercise induced inheritance.

Human exercise sperm promoter DMGs overrepresent targets of exercise related miRNAs

Next, sperm promoter DMGs were subjected to miRNA target enrichment analysis. Targets of various miRNAs were found enriched, with hsa-miR-181 family miRNAs showing highest enrichment (**Fig. 5a**). Notably, exercise associated DEmiRNAs identified in diverse tissues including skeletal muscle, blood cells, and plasma in genome level studies were found inclusive of hsa-miR-181 family miRNAs (**Fig. 5b**). Interestingly, while the latter miRNAs showed upregulation, their targets among the sperm promoter DMGs, namely, CCDC117, GRIK3, IPO8, LHX9, PNRC1, RAB8B, and RFX2, had previously shown mostly hypomethylation and, the corresponding DEmRNAs, wherever present, mostly upregulation (**Fig. 4a**). Because, in the standard paradigm, miRNAs repress gene expression, a negative, not positive miRNA-

mRNA correlation would have been expected. However, at the same time, several instances of positive correlations are also known, with the bidirectionality explained in terms of complex gene regulatory networks (Diaz *et al*., 2015; Laxman *et al*., 2015). Next, gene ontology analysis of hsa-miR-181 family targets showed enrichment of various biological process terms, with several muscle related terms figuring among most highly enriched ones (**Fig. 5c**). These results further consolidated the evidence that exercise sperm DMGs are functionally related to exercise biology.

Human somatic DEmRNAs independently overrepresent targets of exercise related miRNAs

Plasma hsa-miR-181c levels is a predictive marker of exercise training response in heart failure patients (Gevaert *et al*., 2021). Also, circulating levels of hsa-miR-181c in peripheral blood has been associated with exercise capacity and ventilatory inefficiency in acute myocardial infarction (Miyazawa *et al*., 2021). Further, expression levels of miRNAs including hsa-miR-181a in extracellular vesicles in plasma has been correlated with aerobic fitness in young men (Guescini *et al*., 2015). Furthermore, exercise training has been found to cause increased miR-181a expression levels in testicular tissue in opioid treated rats (Hadi *et al*., 2022). Given previous bioinformatic evidence implicating extracellular circulating miRNAs as potential carriers of soma to germline information transfer in epigenetic inheritance (Sharma, 2014), and its subsequent experimental support (Gapp *et al*., 2014; Sharma *et al*., 2016; Chen *et al*., 2016), these hsa-miR-181 family associations in exercise seemed remarkable. Occurrences of hsa-miR-181 family targets were thus examined in exercise and bariatric surgery related DEmRNA genes independent of sperm DMGs (**Fig. 6a**). Strikingly, chi-square goodness of fit test specifically confirmed that the observed counts of the targets were highly significantly more than the expected counts assuming equal frequency of the targets and the non-targets (**Fig. 6b**). To note, positive miRNA-mRNA correlation observed and discussed above was confirmed as hsa-miR-181 family targets also showed upregulation (**Fig. 6a**), suggesting involvement of

complex regulatory networks in exercise response. Overall, these results further amalgamated the evidence that exercise may induce epigenetic inheritance in humans.

The systematic analysis of available omics data presented here compellingly suggests that environmentally induced traits can be inherited in humans. Specifically, physical exercise induced changes in sperm DNA methylome have been found to correlate with exercise induced alterations in coding and non-coding transcriptomes in the soma. The induced changes in germline methylome and somatic transcriptomes also show functional significance in exercise biology. The biological plausibility of exercise induced inheritance in humans is thus strongly supported. Future human cohort studies may provide more definitive evidence. With exercise being a common healthy practice recommended by leading organizations, its effects in the offspring beneficial (Kusuyama *et al*., 2020; Vieira de Sousa Neto *et al*., 2021), sperm and peripheral blood convenient samples, and, now, the biological plausibility for inheritance of exercise induced traits in humans supported, an ethically ideal opportunity has arisen to undertake such studies. In the latter, sperm from prospective fathers in exercise and non-exercise cohorts can be used to confirm expected exercise induced sperm DNA methylation changes, and upon confirmation, the studies can be extended and offspring peripheral blood tested for expected exercise induced gene expression changes. Considering that epigenetic inheritance can offer an explanation for missing causality and heritability of complex diseases (Slatkin, 2009; Furrow *et al*., 2011; Danchin *et al*., 2011; Trerotola *et al*., 2015), and epigenetic drugs a means for reversing disease associated epigenetic states (Heerboth *et al*., 2014; Kronfol *et al*., 2017; Rittiner *et al*., 2022), formal, hypothesis testing human studies on non-DNA sequence-based germline inheritance are warranted. In conclusion, the present human evidence of environmental factor induced gene expression changes in soma and epigenetic changes in germline reflecting each other strongly supports presence of inheritance of acquired traits in our species. The observed equivalence between somatic and germline changes suggests that

given the knowledge of one the other can be predicted. This could be of predictive value in ecology, evolution, health, and disease.

References

1. Adrian-Kalchhauser, I., Sultan, S. E., Shama, L. N. S., Spence-Jones, H., Tiso, S., Valsecchi, C. I. K. & Weissing, F. J. Understanding 'Non-genetic' Inheritance: Insights from Molecular-Evolutionary Crosstalk. *Trends Ecol. Evol.* **35**, 1078-1089 (2020).

2. Agarwal, V., Bell, G. W., Nam, J. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, e05005, (2015).

3. Anastasiadi, D., Venney, C. J., Bernatchez, L. & Wellenreuther, M. Epigenetic inheritance and reproductive mode in plants and animals. *Trends Ecol. Evol.* **36**, 1124-1140 (2021).

4. Anway, M. D., Cupp, A. S., Uzumcu, M. & Skinner, M. K. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* **308**, 1466-1469 (2005).

5. Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, M. & Wishart, D. S. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* **44(W1)**, W147-W153 (2016).

6. Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S. & Soboleva A. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* **41**(Database issue), D991-D995 (2013).

7. Bhalla, A. & Sharma, A. Phenotypic cross-species conservation and cross-generation directionality switching in epigenetic inheritance. *Mol. Ecol.* doi: 10.1111/mec.16969 (2023).

8. Bohacek, J. & Mansuy, I. M. A guide to designing germline-dependent epigenetic inheritance experiments in mammals. *Nat. Methods* **14**, 243-249 (2017).

9. Cavalli, G. & Heard, E. Advances in epigenetics link genetics to the environment and disease. *Nature* **571**, 489-499 (2019).

10. Chen, J., Bardes, E. E., Aronow, B.J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37(Web Server issue)**, W305-11 (2009).

11. Chen, Q, Yan, M., Cao, Z., Li, X., Zhang, Y., Shi, J., Feng, G., Peng, H., Zhang, X., Zhang, Y., Qian, J., Duan, E., Zhai, Q. & Zhou, Q. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* **351**, 397-400 (2016).

12. Danchin, É., Charmantier, A., Champagne, F. A., Mesoudi, A., Pujol, B. & Blanchet, S. Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nat. Rev. Genet.* **12**, 475-486 (2011).

13. Diaz, G., Zamboni, F., Tice, A. & Farci, P. Integrated ordination of miRNA and mRNA expression profiles. *BMC Genomics* **16**, 767 (2015).

14. Donkin, I., Versteyhe, S., Ingerslev, L. R., Qian, K., Mechta, M., Nordkap, L., Mortensen, B., Appel, E. V. R., Jørgensen, N., Kristiansen, V. B., Hansen, T., Workman, C. T., Zierath, J. R. & Barrès, R. Obesity and Bariatric Surgery Drive Epigenetic Variation of Spermatozoa in Humans. *Cell Metab.* **23**, 369-378 (2016).

15. Fitz-James, M. H. & Cavalli, G. Molecular mechanisms of transgenerational epigenetic inheritance. *Nat. Rev. Genet.* **23**, 325-341 (2022).

16. Furrow, R. E. & Freddy, B. Christiansen, and Marcus W. Feldman. Environment-Sensitive Epigenetics and the Heritability of Complex Diseases. *Genetics* **189**, 1377–1387 (2011).

17. Gallego-Paüls, M., Hernández-Ferrer, C., Bustamante, M., Basagaña, X., Barrera-Gómez, J., Lau, C. E., Siskos, A. P., Vives-Usano, M., Ruiz-Arenas, C., Wright, J., Slama, R., Heude, B., Casas, M., Grazuleviciene, R., Chatzi, L., Borràs, E., Sabidó, E., Carracedo, Á., Estivill, X., Urquiza, J., Coen, M., Keun, H. C., González, J. R., Vrijheid, M. & Maitre, L. Variability of multi-omics profiles in a population-based child cohort. *BMC Med.* **19**, 166 (2021).

18. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* **49(D1)**, D325-D334 (2021).

19. Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, J., Prados, J., Farinelli, L., Miska, E. & Mansuy, I. M. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* **17**, 667-669 (2014).

20. Gevaert, A. B., Witvrouwen, I., Van Craenenbroeck, A. H., Van Laere, S. J., Boen, J. R. A., Van de Heyning, C. M., Belyavskiy, E., Mueller, S., Winzer, E., Duvinage, A., Edelmann, F., Beckers, P. J., Heidbuchel, H., Wisløff, U., Pieske, B., Adams, V., Halle, M. & Van Craenenbroeck, E. M. OptimEx-Clin Study Group. miR-181c level predicts response to exercise training in patients with heart failure and preserved ejection fraction: an analysis of the OptimEx-Clin trial. *Eur. J. Prev. Cardiol.* **28**, 1722-1733 (2021).

21. Gramates, L. S., Agapite, J., Attrill, H., Calvi, B. R., Crosby, M., Santos, G. D., Goodman, J. L., Goutte-Gattat, D., Jenkins, V., Kaufman, T., Larkin, A., Matthews, B., Millburn, G., Strelets, V. B. & the FlyBase Consortium. FlyBase: a guided tour of highlighted features. *Genetics* **220**, iyac035 (2022).

22. Guescini, M., Canonico, B., Lucertini, F., Maggio, S., Annibalini, G., Barbieri, E., Luchetti, F., Papa, S. & Stocchi, V. Muscle Releases Alpha-Sarcoglycan Positive Extracellular Vesicles Carrying miRNAs in the Bloodstream. *PLoS One* **10**, e0125094 (2015).

23. Haahr, M. RANDOM.ORG: True Random Number Service. [ONLINE] https://www.random.org (2023).

24. Hadi, G., Asghar, T., Javad, T. A. & Mazdak, R. Exercise training amplifies SIRT1/Nrf2/antioxidant/testosterone pathway after long-time tramadol toxicity in rat testicles; insights into miR-126-3p and miR-181a induced roles. *Biomed. Pharmacother.* **153**, 113332 (2022).

25. Heerboth, S., Lapinska, K., Snyder, N., Leary, M., Rollinson, S. & Sarkar, S. Use of epigenetic drugs in disease: an overview. *Genet. Epigenet.* **6**, 9-19 (2014).

26. Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N. & Mohr, S.E. An integrative approach to ortholog prediction for disease-focused and other functional studies. An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics* **12**, 357 (2011).

27. Ingerslev, L. R., Donkin, I., Fabre, O., Versteyhe, S., Mechta, M., Pattamaprapanont, P., Mortensen, B., Krarup, N. T. & Barrès, R. Endurance training remodels sperm-borne small RNA expression and methylation at neurological gene hotspots. *Clin. Epigenetics* **10**, 12 (2018).

28. Juengst, E. T., Fishman, J. R., McGowan, M. L. & Settersten Jr, R. A. Serving epigenetics before its time. *Trends Genet.* **30**, 427-429 (2014).

29. Kaufman, J., Khan, M., Payne, J. S., Mancini, J. & White, Y. S. Transgenerational Inheritance and Systemic Racism in America. *Psych. Res. Clin. Pract.* doi: 10.1176/appi.prcp.20220043 (2023).

30. Kronfol, M. M., Dozmorov, M. G., Huang, R., Slattum, P. W. & McClay, J. L. The role of epigenomics in personalized medicine. *Expert Rev. Precis. Med. Drug Dev.* **2**, 33-45 (2017).

31. Kusuyama, J., Alves-Wagner, A. B., Makarewicz, N. S. & Goodyear, L. J. Effects of maternal and paternal exercise on offspring metabolism. *Nat. Metab.* **2**, 858-872 (2020).

32. Laxman, N., Rubin, C., Mallmin, H., Nilsson, O., Pastinen, T., Grundberg, E. & Kindmark, A. Global miRNA expression and correlation with mRNA levels in primary human bone cells. *RNA* **21**, 1433–1443 (2015).

33. Lyne, R., Bazaga, A., Butano, D., Contrino, S., Heimbach, J., Hu, F., Kalderimis, A., Lyne, M., Reierskog, K., Stepan, R., Sullivan, J., Wise, A., Yehudi, Y. & Micklem, G. HumanMine: advanced data searching, analysis and cross-species comparison *bioRxiv* doi: 10.1101/2022.05.11.491545 (2022).

34. Lyne, R., Smith, R., Rutherford, K., Wakeling, M., Varley, A., Guillier, F., Janssens, H., Ji, W., Mclaren, P., North, P., Rana, D., Riley, T., Sullivan, J., Watkins, X., Woodbridge, M., Lilley, K., Russell, S., Ashburner, M., Mizuguchi, K. & Micklem, G. FlyMine: an integrated database for Drosophila and Anopheles genomics. *Genome Biol.* **8**, R129 (2007).

35. Mahi, N. A., Najafabadi, M. F., Pilarczyk, M., Kouril, M. & Medvedovic, M. GREIN: An Interactive Web Platform for Re-analyzing GEO RNA-seq Data. *Sci Rep.* **9**, 7580 (2019).

36. Miyazawa, R., Iso, Y., Tsujiuchi, M., Shoji, M., Takahashi, T., Koba, S., Ebato, M., Miyagawa, T., Geshi, E. & Suzuki, H. Potential Association of Circulating MicroRNA-181c and MicroRNA-484 Levels with Cardiorespiratory Fitness after Myocardial Infarction: A Pilot Study. *Prog. Rehabil. Med.* **6**, 20210017 (2021).

37. Molaro, A., Hodges, E., Fang, F., Song, Q., McCombie, W. R., Hannon, G. J. & Smith, A. D. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. *Cell* **146**, 1029-1041 (2011).

38. Motenko, H., Neuhauser, S. B., O'Keefe, M. & Richardson, J. E. MouseMine: a new data warehouse for MGI. *Mamm. Genome* **26**, 325-330 (2015).

39. Nätt, D., Kugelberg, U., Casas, E., Nedstrand, E., Zalavary, S., Henriksson, P., Nijm, C., Jäderquist, J., Sandborg, J., Flinke, E., Ramesh, R., Örkenby, L., Appelkvist, F., Lingg, T., Guzzi, N., Bellodi, C., Löf, M., Vavouri, T. & Anita Öst, A. Human sperm displays rapid responses to diet. *PLoS Biol.* **17**, e3000559 (2019).

40. Niu, Z., Mohazzab-Hosseinian, S. & Breton, C. V. Transgenerational epigenetic inheritance: Perspectives and challenges. *J. Allergy Clin. Immunol.* doi: 10.1016/j.jaci.2023.02.027 (2023).

41. Pepin, M. E., Drakos, S., Ha C. M., Tristani-Firouzi, M., Selzman, C. H., Fang, J. C., Wende, A. R. & Wever-Pinzon, O. DNA methylation reprograms cardiac metabolic gene expression in end-stage human heart failure. *Am. J. Physiol. Heart Circ. Physiol.* **317**, H674-H684 (2019).

42. Perez, M. F. & Lehner, B. Intergenerational and transgenerational epigenetic inheritance in animals. *Nat. Cell Biol.* **21**, 143-151 (2019).

43. Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* **35(Web Server issue)**, W193–W200 (2007).

44. Rittiner, J., Cumaran, M., Malhotra, S. & Kantor, B. Therapeutic modulation of gene expression in the disease state: Treatment strategies and approaches for the development of next-generation of the epigenetic drugs. *Front. Bioeng. Biotechnol.* **17**, 10:1035543 (2022).

45. Santilli, F. & Boskovic, A. Mechanisms of transgenerational epigenetic inheritance: lessons from animal model organisms. *Curr. Opin. Genet. Dev.* **79**, 102024 (2023).

46. Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., Tse, T., Wang, J., Williams, R., Trawick, B. W., Pruitt, K. D. & Sherry, S. T. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **50(D1)**, D20–D26 (2022).

47. Seal, R. L., Braschi, B., Gray, K., Jones, T. E. M., Tweedie, S., Haim-Vilmovsky, L. & Bruford, E. A. Genenames.org: the HGNC resources in 2023. *Nucleic Acids Res.* **51(D1)**, D1003-D1009 (2023).

48. Senaldi, L. & Smith-Raska, M. Evidence for germline non-genetic inheritance of human phenotypes and diseases. *Clin. Epigenetics* **12**, 136 (2020).

49. Sharma, A. Multi-omics data analysis implicating epigenetic inheritance in evolution and disease. *Biochim. Biophys. Acta Gen. Subj.* **1864**, 129477 (2020).

50. Sharma, A. Novel transcriptome data analysis implicates circulating microRNAs in epigenetic inheritance in mammals. *Gene* **538**, 366-372 (2014).

51. Sharma, A. & Singh, P. Detection of transgenerational spermatogenic inheritance of adult male acquired CNS gene expression characteristics using a Drosophila systems model. *PLoS One* **4**, e5763 (2009).

52. Sharma, U., Conine, C. C., Shea, J. M., Boskovic, A., Derr, A. G., Bing, X. Y., Belleannee, C., Kucukural, A., Serra, R. W., Sun, F., Song, L., Carone, B. R., Ricci, E. P., Li, X. Z., Fauquier, L., Moore, M. J., Sullivan, R., Mello, C. C., Garber, M., Rando, O. J. Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* **351**, 391-396 (2016).

53. Silva, G. J. J., Bye, A., El Azzouzi, H. & Wisløff, U. MicroRNAs as Important Regulators of Exercise Adaptation. *Prog. Cardiovasc. Dis.* **60**, 130-151 (2017).

54. Slatkin, M. Epigenetic Inheritance and the Missing Heritability Problem. *Genetics* **182**, 845–850 (2009).

55. Takahashi, Y., Valencia, M. M., Yu, Y., Ouchi, Y., Takahashi, K., Shokhirev, M. Lande, K., Williams, A. E., Fresia, C., Kurita, M., Hishida, T., Shojima, K., Hatanaka, F., Nuñez-Delicado, E., Esteban, C. R. & Belmonte, J. C. I. Transgenerational inheritance of acquired epigenetic signatures at CpG islands in mice. *Cell* **186**, 715-731.e19 (2023).

56. Trerotola, M., Relli, V., Simeone, P. & Alberti, S. Epigenetic inheritance and the missing heritability. *Hum. Genomics* **9**, 17 (2015).

57. Vågerö, D., Cederström, A. & van den Berg, G. J. Food abundance in men before puberty predicts a range of cancers in grandsons. *Nat. Commun.* **13**, 7507 (2022).

58. Vieira de Sousa Neto, I., Fontes, W., Prestes, J. & de Cassia Marqueti, R. Impact of paternal exercise on physiological systems in the offspring. *Acta Physiol. (Oxf)* **231**, e13620 (2021).

59. Vina, J., Sanchis-Gomar, F., Martinez-Bello, V. & Gomez-Cabrera, M. C. Exercise acts as a drug; the pharmacological benefits of exercise. *Br. J. Pharmacol.* **167**, 1-12 (2012).

60. Wang, K., Li, M. & Hakonarson, H. Analysing biological pathways in genome-wide association studies. *Nat. Rev. Genet.* **11**, 843-854 (2010).

61. Wang, Z., Neuburg, D., Li, C., Su, L., Kim, J. Y., Chen, J. C. & Christiani, D. C. Global gene expression profiling in whole-blood samples from individuals exposed to metal fumes. *Environ. Health Perspect.* **113**, 233-241 (2005).

62. Watanabe, L. P. & Riddle, N. C. New opportunities: Drosophila as a model system for exercise research. *J. Appl. Physiol. (1985)* **127**, 482-490 (2019).

Conflict of interest

The author declares no conflict of interest

Figure legends

Fig. 1. Gene set overlap between human sperm DMGs and somatic DEmRNAs. Heatmaps showing methylation difference in exercise (**a**) and bariatric surgery (**b**) associated sperm DMGs at two reported time points, and showing fold change of most common DEmRNAs in skeletal muscle (**c**) and blood (**d**) following exercise, and in (**e**) subcutaneous adipose tissue following bariatric surgery. The number of DMGs and DEmRNAs in the heatmaps are indicated. b, blood; bs, bariatric surgery; ex, differentially expression; exr, exercise; FC, fold change; h, human; leuk, leukocytes; MD, methylation difference; me, differential methylation; pbmc, peripheral blood mononuclear cells; sat, subcutaneous adipose tissue; sm, skeletal muscle; sp, sperm; wb, whole blood. Individual samples are intuitively labeled, with intervention, species, tissue, data type, and sample category multiplicity indicated in that order and separated by a dash. (**f**) Bar graph showing fold enrichment of pooled sets of skeletal muscle (exr-h-sm-ex), blood (exr-h-b-ex), and subcutaneous adipose tissue (bs-h-sat-ex) DEmRNAs in the pooled set of exercise associated sperm DMGs (exr-h-sp-me) and size-matched non-overlapping subsets of bariatric surgery associated sperm DMGs (bs-h-sp-me_i to _x). FE, fold enrichment.

Fig. 2. Gene set overlap between human sperm DMGs and animal model somatic DEmRNAs. Heatmaps showing fold change of most common exercise associated DEmRNAs in rat (**a**) and mouse (**b**) skeletal muscle, and of (**c**) exercise associated DEmRNAs in both larval body wall muscles and whole organism fly samples. r, rat; m, mouse; f, fly; bwm, larval body wall muscles; wo, whole organism. Labeling scheme of individual samples, and abbreviations as explained in Figure 1. (**d**) Bar graph showing fold enrichment of pooled sets of human orthologs of rat (exr-r-sm-ex) and mouse (exr-m-sm-ex) skeletal muscle, and of larval body wall muscles and whole fly (exr-f-ex) DEmRNAs in the pooled sets of exercise and bariatric surgery associated human sperm DMGs. Other details as mentioned in Figure 1.

Fig. 3. Gene ontology enrichment in human and animal model DMGs and/or DEmRNAs. (**a**) Heatmap showing fold enrichment of biological process terms most commonly overrepresented among pooled sets of exercise associated human sperm DMGs (exr-h-sp-me), human skeletal muscle (exr-h-sm-ex) and blood (exr-h-b-ex) DEmRNAs, rat (exr-r-sm-ex) and mouse (exr-m-sm-ex) skeletal muscle DEmRNAs, and *Drosophila* (exr-f-ex) DEmRNAs. (**b**) Heatmap showing fold enrichment of all "muscle" containing terms. The full list of enriched terms is provided in **Supplementary Table 2**. Other details as mentioned in Figure 1 and 2.

Fig. 4. Directionality of human sperm promoter DMGs and somatic tissue DEmRNAs. (**a**) Heatmap of exercise associated human sperm promoter- and non-promoter DMGs (methylation difference) that are common for both the reported time points, shown along with corresponding DEmRNAs (fold change) in human exercise associated skeletal muscle and blood, and bariatric surgery associated subcutaneous adipose tissue. DEmRNA samples present in Figure 1C-E but missing here indicate no DMG common DEmRNA therein. A reduced number of heatmap shades was used to discretize regulation. (**b**) Bar graph showing expected and observed total counts of DEmRNA occurrences with regulatory direction inverse of

DMGs. Chi-square goodness of fit and corresponding p values are shown above the bars. Other details as mentioned in Figure 1.

Fig. 5. Enrichment of miRNA targets in human sperm promoter DMGs. (**a**) Bar graph showing Bonferroni *q* value of fold enrichment for targets of various miRNAs. The dashed line indicates cutoff for statistical significance. (**b**) Heatmap showing fold change of DEmiRNAs most commonly identified among various tissues in human volunteers following exercise. The hsa-miR-181 family miRNAs are blown-up in the inset. neut, neutrophils; nk, natural killer cells; pl, plasma. (**c**) Bar graph showing fold enrichment of most overrepresented biological process terms in hsa-miR-181 family targets. The complete list of enrichment terms is given in **Supplementary Table 3**. Other details as mentioned in Figure 1.

Fig. 6. Independent enrichment of miRNA targets in human somatic DEmRNAs. (**a**) Heatmap showing fold change of human exercise associated DEmRNAs in skeletal muscle and blood, and bariatric surgery associated DEmRNAs in subcutaneous adipose tissue. The genes in the exercise associated DEmRNAs are exclusive of exercise sperm promoter DMGs. Samples present in Figure 1C-E but missing here indicate no hsa-miR-181 family target therein. A reduced number of heatmap shades was used to discretize regulation. (**b**) Bar graph showing expected and observed total counts of hsa-miR-181 family targets. Chi-square goodness of fit and corresponding p values are shown above the bars. Other details as mentioned in Figure 1.

Supplementary materials

Supplementary Table 1. Data compendium

Supplementary Table 2. Gene ontology enrichment in exercise associated gene sets

Supplementary Table 3. Gene ontology enrichment in has-miR-181 family targets

Fig. 1

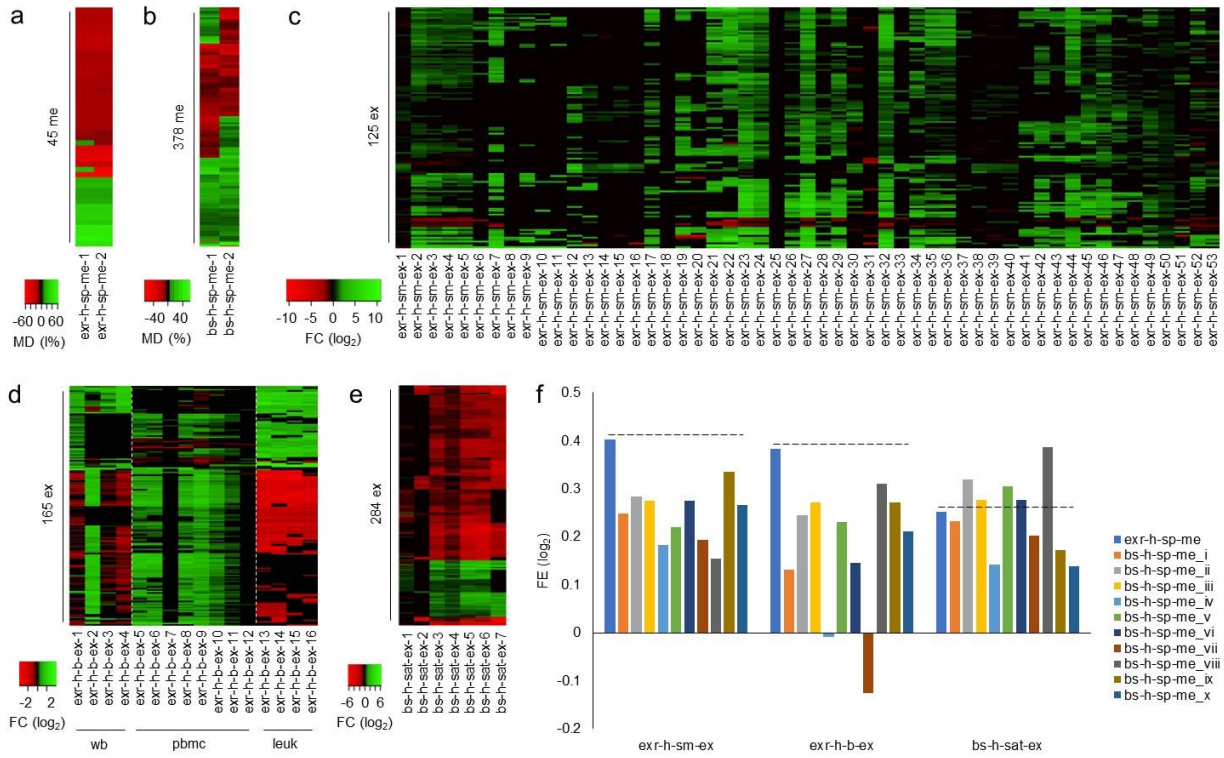
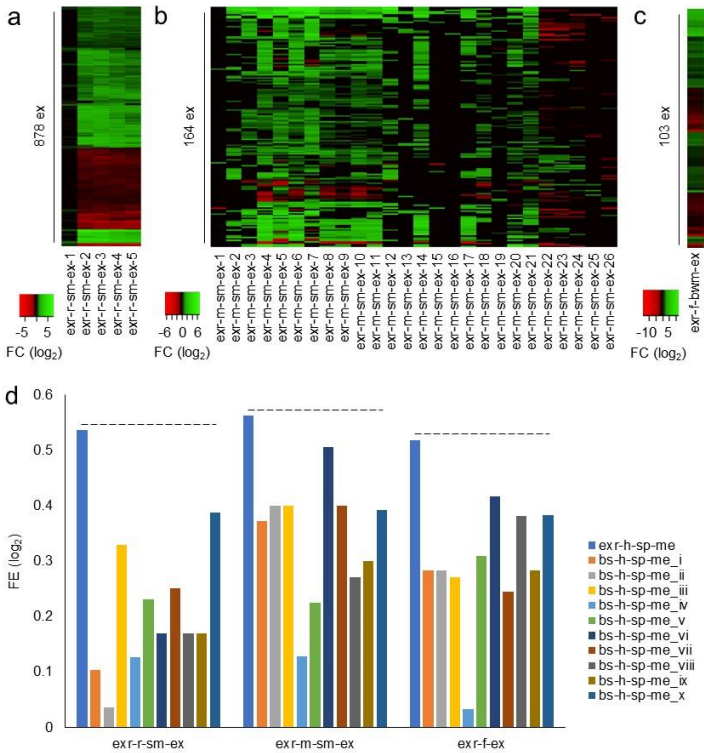
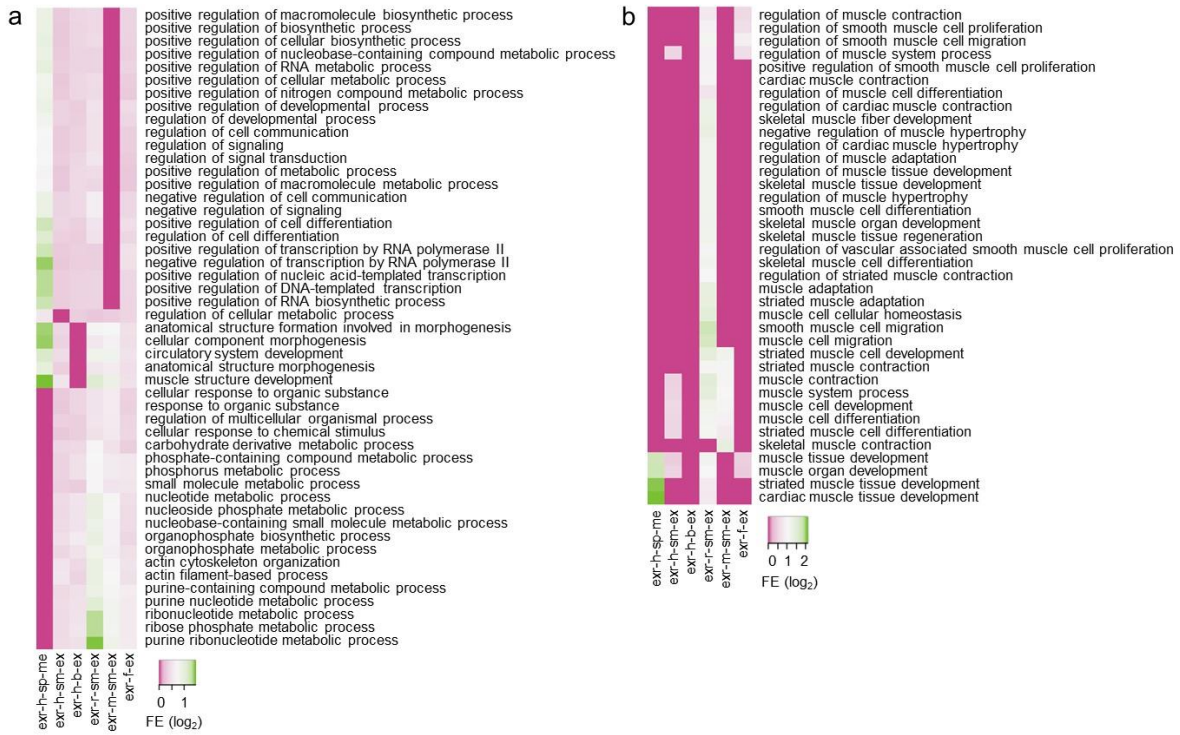

a
45 me
exr-h-sp-me-1
exr-h-sp-me-2
-60 0 60
MD (l%)
b
378 me
bs-h-sp-me-1
bs-h-sp-me-2
-40 40
MD (%)
c
125 ex
-10 -5 0 5 10
FC (log2)
d
165 ex
-2 2
FC (log2)
wb
pbmc
leuk
e
284 ex
-6 0 6
FC (log2)
f
FE (log2)
exr-h-sm-ex
exr-h-b-ex
bs-h-sat-ex
exr-h-sp-me
bs-h-sp-me_i
bs-h-sp-me_ii
bs-h-sp-me_iii
bs-h-sp-me_iv
bs-h-sp-me_v
bs-h-sp-me_vi
bs-h-sp-me_vii
bs-h-sp-me_viii
bs-h-sp-me_ix
bs-h-sp-me_x

Fig. 2


a
878 ex
exr-r-sm-ex-1
exr-r-sm-ex-2
exr-r-sm-ex-3
exr-r-sm-ex-4
exr-r-sm-ex-5
-5 5
FC (log2)
b
164 ex
exr-m-sm-ex-1
exr-m-sm-ex-2
exr-m-sm-ex-3
exr-m-sm-ex-4
exr-m-sm-ex-5
exr-m-sm-ex-6
exr-m-sm-ex-7
exr-m-sm-ex-8
exr-m-sm-ex-9
exr-m-sm-ex-10
exr-m-sm-ex-11
exr-m-sm-ex-12
exr-m-sm-ex-13
exr-m-sm-ex-14
exr-m-sm-ex-15
exr-m-sm-ex-16
exr-m-sm-ex-17
exr-m-sm-ex-18
exr-m-sm-ex-19
exr-m-sm-ex-20
exr-m-sm-ex-21
exr-m-sm-ex-22
exr-m-sm-ex-23
exr-m-sm-ex-24
exr-m-sm-ex-25
exr-m-sm-ex-26
-6 0 6
FC (log2)
c
103 ex
exr-f-bwm-ex
exr-f-wo-ex
-10 5
FC (log2)
d
FE (log2)
0.6
0.5
0.4
0.3
0.2
0.1
0
exr-r-sm-ex
exr-m-sm-ex
exr-f-ex
exr-h-sp-me
bs-h-sp-me_i
bs-h-sp-me_ii
bs-h-sp-me_iii
bs-h-sp-me_iv
bs-h-sp-me_v
bs-h-sp-me_vi
bs-h-sp-me_vii
bs-h-sp-me_viii
bs-h-sp-me_ix
bs-h-sp-me_x

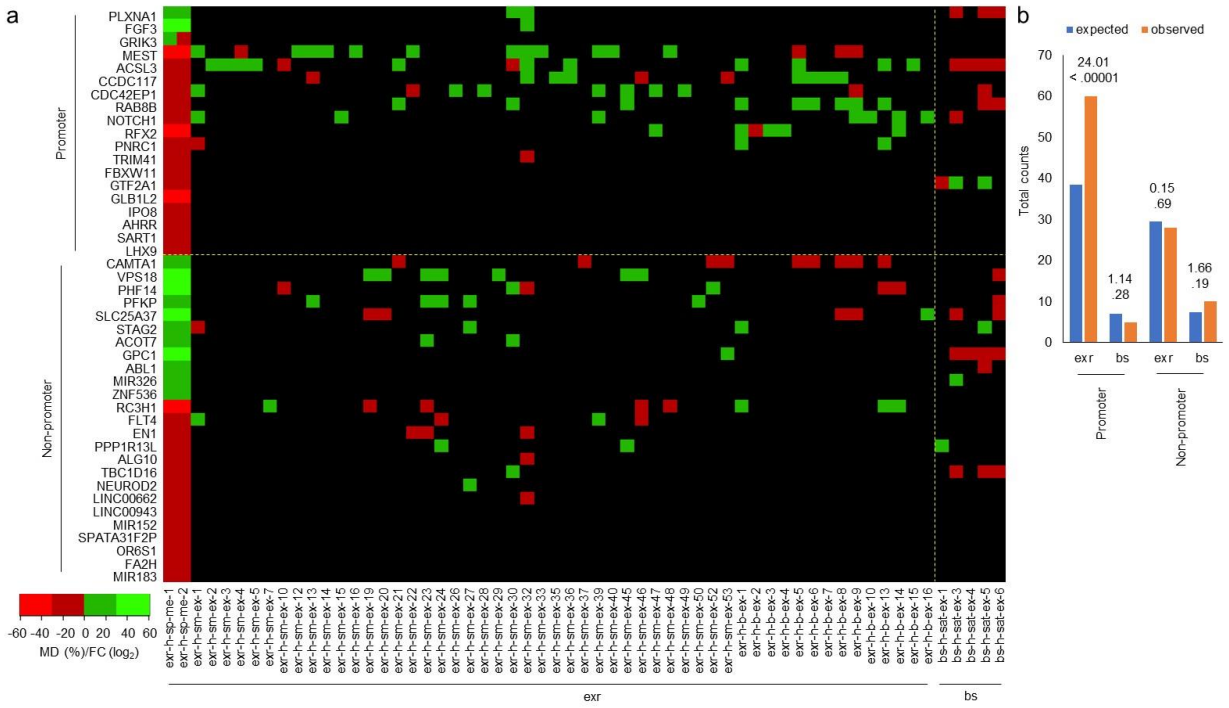
580 **Fig. 3**



581

582

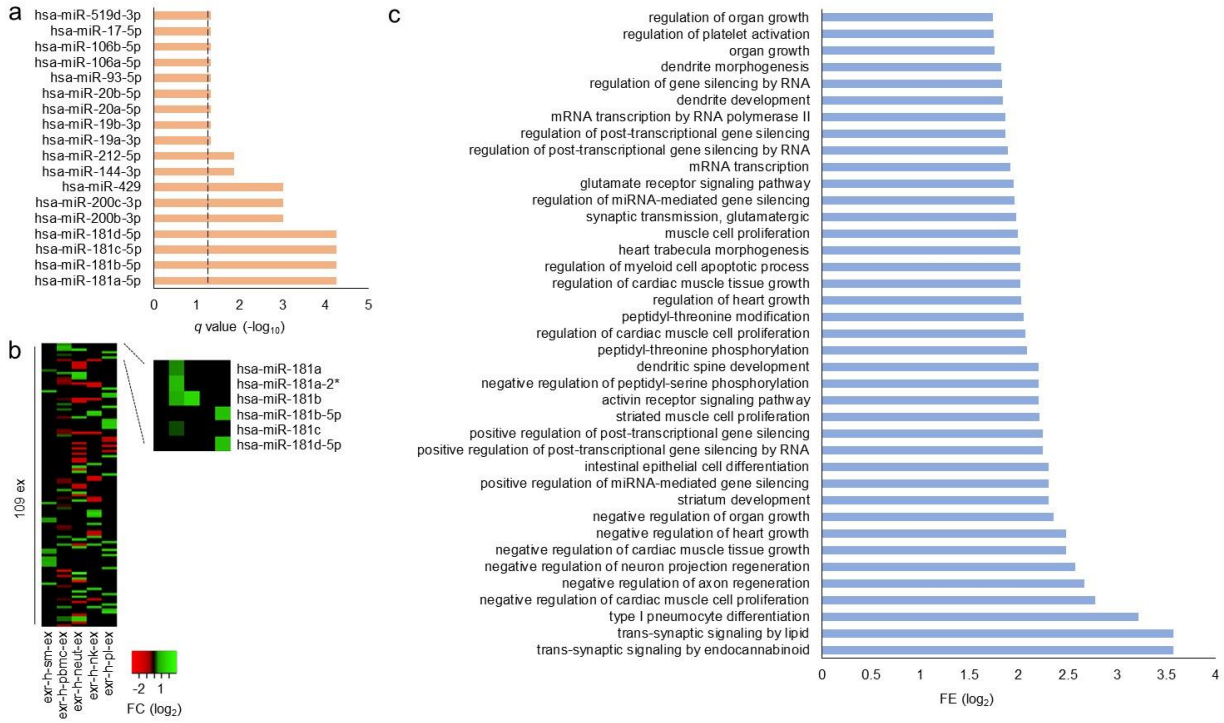
583 **Fig. 4**



584

585

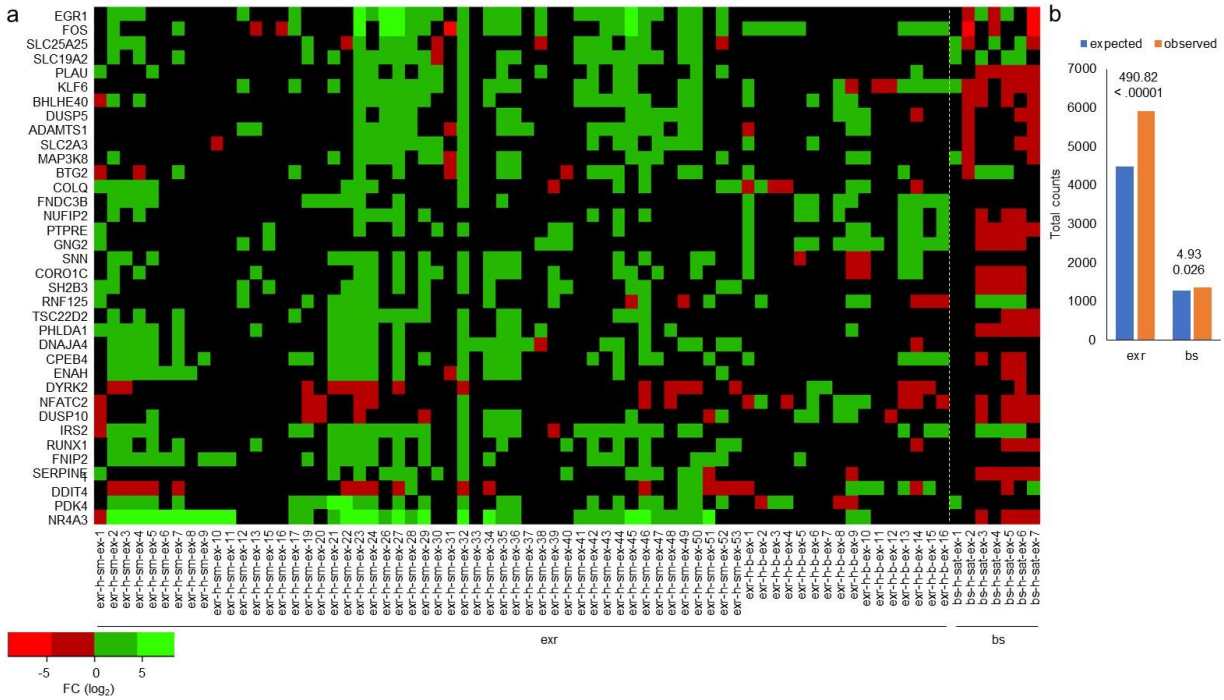
586 **Fig. 5**



587

588

589 **Fig. 6**



590

591

592

593