A role for heterochromatin and repetitive elements in epigenetic inheritance

Eduard Casas Masnou

A role for heterochromatin and transposable elements in epigenetic inheritance

Eduard Casas Masnou

Programa Doctoral en Biomedicina, UB, 2019

Institut de Recerca Germans Trias i Pujol (IGTP), Badalona, Catalunya, Spain

Thesis director

Dr. Tanya Vavouri

Regulatory Genomics, Institut de Recerca Germans Trias i Pujol (IGTP), Badalona, Catalunya, Spain

Josep Carreras Leukaemia Research Institute (IJC), Badalona, Catalunya, Spain

Thesis tutor

Dr. Modesto Orozco

Department of Biochemistry and Biomedicine, UB, Barcelona, Catalunya, Spain

Acknowledgments

Vull començar aquesta tesi agraïnt a tota la meva gent el suport i l'ajuda sense els quals no hagués estat possible realitzar-la. Aquesta tesi doctoral és el final d'un projecte de desenvolupament personal, alhora que científic, que per a mi va començar fa molts més anys. He tingut la immensa sort de compartir aquests anys amb gent magnífica que m'ha acompanyat al llarg de tot aquest temps.

Començaré per la meva família, en Jordi, la mama i en Guillem. M'han ajudat a desenvolupar la meva curiositat científica ja que, des de petit, sempre han recolzat les meves inquietuds i aptituds acadèmiques. Tot això apart de contribuir substancialment a la logística, la intendència i al sentit comú del meu manteniment al llarg d'aquests 30 anys.

Vull agraïr a la meva colla, els companys d'escalada i de pis, l'Albert, en Pau, en Nico, en Lluís, en Juan, en Petr, i a la Maria, per tot el temps plegats fent aquests anys molt més interessants i emocionants.

A la meva segona família, la gent de Veterinària, en Nil, en Lluís, l'Emili, l'Oriol, en Martí, en Marc, en Josep i la Júlia, que, tot i que la majoria ja estiguin ara casats i amb fills, els segueixo estimant molt.

També als amics de Girona, especialment l'Arnald i en Sergi, amb qui tot i que passin els anys sempre ens sentim aprop.

Finalment, es mereixen també un agraïment especial la gent del Campus amb qui hem passat tantes hores plegats dins i fora de la feina. Gent amb qui hem anat a còrrer, caminar, esquiar, jugar a volei, a jocs de taula, a menjar i beure, i a mil coses més. En concret penso amb en Joan, en Marc, la Laura, la Marta i l'Anna que m'han aportat un últim any d'escritura de tesi genial, i amb la gent més 'senior' amb qui hem compartit els dinars a la feina, en Roberto i l'Emili, i d'altres amics que han seguit la carrera a altres llocs com l'Iñaki i la Izaskun.

A tots i totes els vull fer arribar el meu reconeixement i agraïment per haver-me acompanyat i encoratjat a arribar fins aquí.

I also want to acknowledge all collaborators whose work made this thesis possible, Josep, Sonia and Cristina and people with whom I worked very closely these years like Adam, Anita, Heidi, Ben and especially Andrew.

Last but not least, obviously, I want to give big thanks to Tanya, for being an excellent PhD supervisor, always guiding, helping, always presenting new ideas when I was stuck and being supportive with my nonsense ideas. Moreover, I want to especially thank you for being exceptionally great at managing the lab in terms of happiness and enjoyment and realization at work. I would say that the big success of my PhD is having had a great time throughout these six years and nothing shouts it out louder that saying that I would start over again.

Summary

Gene regulation mechanisms control the level of transcription of each gene into RNA and the combination of expressed genes determines cell identity. Gene regulation is maintained by epigenetic mechanisms including DNA methylation, histone modifications and non-coding RNAs. These same mechanisms are responsible for silencing of transposable elements and heterochromatin formation. Interestingly, epigenetic mechanisms can transmit the transcriptional state of a gene to the next generation. Epigenetic inheritance differs from conventional genetics: it does not follow the law of segregation and importantly, can transmit acquired traits.

The advent of next generation sequencing (NGS) allows for gene expression quantification and epigenome profiling, opening the door to genome wide screenings of epigenetic factors and phenotypes linked to epigenetic inheritance. Here, I study the role of heterochromatin and transposable elements in epigenetic inheritance.

In the first chapter I present how an IAP insertion in the Nocturnin gene triggers the birth of new piRNA cluster in mouse. We hypothesize that many piRNA producing loci have evolved from ERV insertions into germline expressed genes. Last, we identify NXF1 as a key factor in piRNA biogenesis of IAP-derived piRNA loci.

In the second chapter I test whether the IAP insertion in Nocturnin, and therefore piRNAs produced from this gene in the male germline, affect expression of the gene in the embryo. I find that the piRNA-producing allele of Nocturnin is more highly expressed from the paternal that the maternal allele in early embryo. Thus, the IAP insertion in Nocturnin leads to transmission of an altered epigenetic expression state from parents to progeny, potentially via the production of piRNAs in the male germline.

In the third chapter of this thesis I describe a model of intergenerational epigenetic inheritance in flies. My work describes genome wide changes in gene expression that are direct consequences of epigenetic inheritance and I identify chromatin factors related to the transmission and maintenance of the phenotype in the next generation.

In the fourth chapter of my thesis I use worms exposed to high temperature to identify endogenous genes that are able to maintain memory of expression for many generations. Interestingly, I find that transposable elements that are upregulated by temperature and repressed by heterochromatin can transmit epigenetic information to the progeny.

In the fifth chapter of this thesis I study heritable expression of acquired expression states after epigenetic information loss linked to impaired DNA replication. My work describes how the loss of repressive marks during impaired replication in embryos leads to heritable changes in gene expression of loci regulated by heterochromatin and polycomb means.

Index

Introduction

On Genetics and inheritance

The study of heredity has attracted the attention of humankind from early ages. Already in the 19th century Mendel described the principles behind basic genetics with the laws of segregation (Mendel 1865). Phenotypic information responsible for the characterization of an organism is stored within a continuous sequence of nucleotides of deoxyribonucleic acid (DNA) structured in a characteristic double-helix structure (Watson and Crick 1953).

The whole genetic information stored in DNA is called the genome and is the blueprint of all known organisms. The first human genome was sequenced in the 21st century (Lander et al. 2001). The main discrete, functional units of the genome are called genes (Mendel 1865). Genes are transcribed into RNA and RNA is translated into proteins (Crick 1970) according to the genetic code (Nirenberg and Matthaei 1961). However, genes, and specifically protein-coding genes, make up only 1% of the human genome, and we know very little about the functionality of the other 99% of the genome.

In sexual reproduction the genome of each parent is copied to gametes. Gametes (hence genomes) of each sex fuse into a zygote which features mutations and recombinations of the parental genomes that make each organism genetically unique. This turns into phenotypic diversity which allows organisms to increase their fitness and species to adapt (Darwin 1859).

Yet, during the 20th century a parallel form of inheritance has been described: the transmission of acquired traits (Waddington 1953). This paradigm of inheritance different to conventional genetics is not gene-based and does not follow the law of segregation. This type of inheritance is called epigenetic (see below). Many attempts have been made to disentangle the mechanisms behind epigenetic inheritance in animals (Perez and Lehner 2019). Still, there are many unknowns (reviewed in (Heard and Martienssen 2014).

The importance of gene regulation

All cells of the same multicellular organism develop from the same zygote hence share copies of the same genome. However, all cell lineages behave differently. This is in part because genes are expressed -transcribed- at different levels in different cell types. Gene regulation mechanisms control the level of transcription of each gene into RNA. The specific combination of expressed genes determines cell identity (reviewed in (Eckersley-Maslin, Alda-Catalinas, and Reik 2018). Gene regulation grants one of the most amazing phenomena in eukaryotes: the development from a single zygotic cell into hundreds of highly specialized cell types.

Development is driven mainly by transcription factors that bind specific DNA sequences -transcription factor binding sites or motifs- and regulate the expression of lineage-specific genes (reviewed in (Srivastava and DeWitt 2016). However, a specific expression pattern needs to be maintained through time after the initial trigger. The mechanisms by which cells maintain memory of expression after the initial stimulus is withdrawn are generally called epigenetics.

Epigenetics: mechanisms of genome regulation

The term epigenetics was coined by the geneticist and developmental biologist Conrad H. Waddington as the study of the mechanisms that influence the phenotype given the genotype (Waddington 1942), albeit the definition has largely evolved in recent years. Nowadays we refer to epigenetics as the mechanisms that maintain memory of a phenotype without a change in DNA sequence (Deans and Maggert 2015).

Epigenetic mechanisms include DNA methylation, histone modifications and non-coding RNA among others. Epigenetic states are, for example, responsible for the maintenance of gene regulation, hence cell identity, and also for the repression of transposons. These states are stably maintained through time and importantly, inherited after cell division.

Epigenetic states can be influenced by external or environmental stimuli too, modulating gene expression thus, phenotypes. For instance, temperature experienced during embryonic development determines the sex in many reptiles (Charnier 1966; Bull and Vogt 1979; Ferguson and Joanen 1982) and fish (reviewed in (Ospina-Alvarez and Piferrer 2008) through epigenetics.

Epigenetic states can also be inherited by the next generation. Observations of inheritance of epigenetic expression states not following mendelian laws drew the attention of scientists many years ago (McClintock and Others 1958; McClintock 1961). Interestingly, most of the genes that are able to transmit memory of expression are genes regulated by transposable elements (reviewed in (Slotkin and Martienssen 2007). We call epialleles the heritable, alternative expression states of a gene not caused by genetic variation (Rakyan et al. 2002). This epigenetic state can be transmitted to the next generation through the gametes.

Importantly, the germline epigenome has to display great plasticity so that environmental perturbations can affect it and it can transmit acquired phenotypes to the next generation. In contrast to Weisman's hypothesis (who suggested that parents only transmit to offspring the hereditary material from germ cells, not information from somatic or

body cells) (Weismann 1893), nowadays there is evidence that plasticity comes in part from soma-to-germline communication that affects the pool of small RNAs in germ cells (Sharma et al. 2018; Posner et al. 2019). This opens the door to transmission of inheritance of gene regulation through small RNAs. Furthermore, there is evidence that the very same germline nucleosomes with histone modifications can be inherited (Gaydos, Wang, and Strome 2014; van der Heijden et al. 2006), making them alternative carriers of epigenetic information. Overall, transgenerational epigenetic inheritance (TEI) is a complex process that can work on multiple pathways, potentially overlapping, so hypotheses of small RNA- and chromatin-based epigenetic inheritance are both feasible. This is visited in far more detail in the following chapters in the introduction.

Yet, epigenetic information encoded in germline has to endure genome-wide reprogramming. Repressive histone modifications and DNA methylation are largely erased (reviewed in (Reik, Dean, and Walter 2001). Actually, epigenetic reprogramming is the mechanism that ensures that histone modification and DNA methylation patterns are mostly erased and not transmitted to the next generation. It is required so that two gametes, highly differentiated cells, can give rise to a totipotent embryo (reviewed in (Reik, Dean, and Walter 2001; Sasaki and Matsui 2008).

Chromatin, heterochromatin and histone modifications

DNA has to be highly condensed to fit into a cell nucleus. To achieve high levels of compaction it is wrapped around proteins into a complex called chromatin. Chromatin is an architectural structure that works as a storage solution but is also key for gene regulation. Chromatin can present different forms, ranging from highly open to highly compacted, but in general it is separated into two major states: euchromatin is highly accessible to transcription polymerases and includes the majority of genes while heterochromatin is highly compacted and generally associated with repetitive elements and repressive states.

Chromatin is structured in nucleosomes around which DNA is wrapped. Nucleosomes are composed of 2 copies of H2A, H2B, H3 and H4 histone proteins (Luger et al. 1997). Histones can be post-translationally modified, especifically on lysines residues of histone H3, to confer gene expression and functional properties to genes in the DNA sequence wrapped around them. Hence, chromatin is a major factor in gene regulation as a carrier of epigenetic information.

Histone modifications contribute to developmental gene regulation by restricting and promoting lineage-specific gene expression. Importantly, chromatin modifications can be stably maintained after cell division and in long periods of time. Histones modifications include methylation and acetylation (among others) and correlate with enhancer activity (H3K4me1) (Heintzman et al. 2009) active promoters (H3K4me3) (Hon, Hawkins, and Ren 2009), transcription (H3K27ac and H3K36me3) (Kolasinska-Zwierz et al. 2009), or silencing (H3K27me3 and H3K9me3) (Z. Wang et al. 2008) among others. Propagating active and silent activity states through time contributes to maintain cell identity.

Chromatin carries epigenetic information through mitosis in many biological processes. For instance in *Drosophila melanogaster* position effect variegation (PEV) is the change in gene expression depending on the genomic context of the gene. Variegation refers to stochastic expression of a gene in different cells of the same tissue resulting in a mosaic pattern and PEV refers to the variegated expression depending

on chromosomal location, usually through abnormal juxtaposition to heterochromatin (reviewed in (Elgin and Reuter 2013)). The phenomenon was first described in 1930 (Muller 1930) in *D. melanogaster*. It was later discovered that chromosomal rearrangements can lead to stochastic epigenetic silencing of genes by spreading of heterochromatin (Schultz 1936). Mutational screenings of genes disrupting PEV allowed the characterization of proteins and histone modifications associated with heterochromatin silencing. A commonly used reporter is the eye-pigment gene *white*, which affects the fly's eyes' color intensity depending on its epigenetic state. Through an artificial vector, *white* is located flanking Fab-7, a Polycomb-regulated enhancer. Polycomb spreads from Fab-7 to *white*, silencing the gene and leading to pale, white eyes (Zink and Paro 1995), suggesting repression by these proteins and the role of epigenetics in maintaining gene repression. Such inactive epigenetic state can be reactivated by an embryonic pulse of transcription, leading to red eyes again (G. Cavalli and Paro 1998). Activation is associated with H4 hyperacetylation which points at this histone modification as the epigenetic mechanism responsible for the maintenance of the expression state (G. Cavalli and Paro 1999).

Chromatin is also linked to temperature-dependent sex determination in *Trachemys scripta elegans* (Ge et al. 2018). KDM6B is a temperature-sensitive histone demethylase that eliminates the repressive mark H3K27me3 near the promoter of *Dmrt1* -the gene that determines male sex- causing its transcription (Ge et al. 2018)). This makes some embryos develop male or female gonads depending on the growth temperature during embryogenesis.

Of specific interest to us is heterochromatin, that is organized in transcriptionally repressed domains (reviewed in (Allshire and Madhani 2018). The repressing capacity of heterochromatin is key to restrain transcription of mobile elements (Rowe et al. 2010). Heterochromatin can be divided into constitutive and facultative heterochromatin (Brown 1966). Constitutive heterochromatin is mainly characterized by H3K9 methylation. Histone methyl-transferases that catalyze mono-, di and tri- methylation of H3K9 include *suppressor of variegation 3-9* (SUV39) in flies and mammals and SET-25 in *C. elegans.* It is usually featured in centromeric or

pericentric domains. Facultative heterochromatin includes more dynamic silencing complexes feature by methylation of H3K27. H3K27me is deposited by the Polycomb repressive com-plex 2 (PRC2) which recruits a different Polycomb (PRC1). Polycomb repression is the main mechanism responsible for the X chromosome inactivation in female mammals (J. Wang et al. 2001).

DNA methylation and genomic imprinting

DNA methylation, more specifically 5-methylcytosine, is nonexistent in *C. elegans* (Simpson, Johnson, and Hammen 1986) and *D. melanogaster (Capuano et al. 2014)*. DNA methylation acts as another layer of genome regulation. It is the main DNA modification and it is typical of CpG dinucleotides. The addition of the methyl group is catalyzed by DNA methyl-transferases (DNMTs) and can be separated in maintenance of DNA methylation after replication by DNMT1 and de-novo DNA methylation by DNMT3A and DNMT3B (Okano et al. 1999).

DNA methylation is key to silence repeats and protect against transposition (Yoder, Walsh, and Bestor 1997; Walsh, Chaillet, and Bestor 1998). However, it also plays a role in regulating gene expression. Gene promoters are enriched in CpG dinucleotides clusters called CpG islands. In general, 5mC of CpG islands overlapping promoters is associated with transcription repression, while 5mC is absent in promoters of transcriptionally active genes.

The most stable pattern of DNA methylation in gene repression is genomic imprinting (reviewed in (Tucci et al. 2019): a mechanism responsible for silencing of genes in a parent-of-origin specific manner. Imprinting is established in the germline and is required for development and stable during adult life (Barton, Surani, and Norris 1984; Surani, Barton, and Norris 1984; McGrath and Solter 1984). Imprinted loci are generally regulated by an Imprinting Control Regions (ICR). These ICRs retain DNA methylation during the post-fertilization epigenetic reprogramming. To note, it was recently described a type of genomic imprinting independent of DNA methylation and based on H3K27me3 (Inoue et al. 2017). Even though there is a lot known about the epigenetic basis of imprinting, there is little known about how species evolve to acquire imprinting at a specific gene to switch off the expression of one parental allele.

Imprinted genes regulate fetal growth and brain function, and disorders associated with imprinting misregulation are responsible for many diseases in humans (reviewed in (Butler 2009).

The PIWI pathway and the role of piRNAs

TE repression is crucial in germ cells, where mutagenic TE insertions can affect the viability of both germline and offspring. In addition, germ cells undergo erasure and resetting of DNA methylation and repressive chromatin creating a window of opportunity for TE activation. This vulnerability is overcome by the germline specific PIWI/piRNA pathway (Carmell et al. 2007; A. A. Aravin et al. 2007, 2008; De Fazio et al. 2011; Reuter et al. 2011; Di Giacomo et al. 2013).

The PIWI pathway is in charge of transcriptional and post transcriptional repression of transposons in mammalian germline (reviewed in (Czech et al. 2018) and is a key player in flies and nematodes, where DNA methylation is absent. How the PIWI pathway responds to the fast changing transposon landscape is largely unknown. It is also required for spermatogenesis (Deng and Lin 2002; Kuramochi-Miyagawa et al. 2004; Carmell et al. 2007) and recently, PIWI proteins have also been involved in maintenance of TEI (Grentzinger et al. 2012; Moore, Kaletsky, and Murphy 2019; Brennecke et al. 2008).

PIWI proteins are members of the Argonaute family, proteins expressed specifically in germ cells (reviewed in (Ernst, Odom, and Kutter 2017)). In mouse there are three *Piwi* family members: *Piwil1* (also known as *Miwi*), *Piwil2* (also known as *Mili),* and *Piwil4 (*also known as *Miwi2)*. MILI and MIWI2 are expressed in the cytoplasm and nucleus of germ cells respectively during embryogenesis and silence repetitive elements transcriptionally (Kuramochi-Miyagawa et al. 2008; A. A. Aravin et al. 2008; De Fazio et al. 2011; Carmell et al. 2007). MILI is expressed throughout spermatogenesis. MILI together with MIWI targets TE transcripts for degradation and also regulates the spermatogenic expression program (Reuter et al. 2011; Vourekas et al. 2012).

PIWI proteins bind a class of small noncoding RNAs called PIWI-interacting RNAs (piRNAs). piRNAs are transcribed from long precursors known as piRNA clusters that are then processed into multiple 26-31 nucleotide long small RNAs (Girard et al. 2006; A. Aravin et al. 2006; Grivna et al. 2006; Brennecke et al. 2007; Lau et al.

2006). Both protein-coding genes and long non-coding RNAs act as piRNA precursor transcripts. After transcription, precursor transcripts are exported to the cytoplasm, where they get processed into mature piRNAs and loaded onto PIWI proteins (Han et al. 2015; Mohn, Handler, and Brennecke 2015). These primary piRNAs, also called phased or trailing piRNAs, often start with uridine. Phased slicing is a mechanism of spreading piRNA sequences and diversity from an existing RNA (Fig 1A).

Secondary piRNAs are generated after initial cleavage of piRNA precursors through the ping-pong cycle (Gunawardane et al. 2007; Brennecke et al. 2007) (Fig 1B). An initiator piRNA guide PIWI proteins and target RNAs by sequence complementarity, cleaving the target transcript at the 10th nucleotide after the first nucleotide of the targeting piRNA generating a responder piRNA. The new responder piRNA is then loaded into PIWI proteins and can act as an initiator closing the ping-pong cycle of piRNA biogenesis. These piRNAs often have an adesonie at position 10 given the reverse complementarity with a uridine at the first position of the initiator piRNAs. The ping pong amplification increases the abundance of the same piRNAs by targeting back identical molecules to the original ones from which piRNAs were generated.

There are two groups of piRNAs in mice depending on the stage at which they are expressed. piRNAs expressed before the pachytene stage of meiosis (pre-pachytene piRNAs) include piRNAs that target transposons and piRNAs spanning the coding sequence and the 3' UTR of many genes. Pachytene piRNAs are generated after the transcription factor A-MYB binds the ~100 piRNA clusters promoting transcription of these piRNA's precursors. Importantly, it is not known how transcripts of protein-coding genes are selected for piRNA production rather than mRNA translation (reviewed in (Ozata et al. 2019) nor we have a clue about the role of these piRNAs.

piRNAs repress transposons by guiding PIWI proteins to target RNAs by sequence complementarity (Fig 1C). piRNAs also promote transposon silencing through heterochromatin assembly in flies (Le Thomas et al. 2014) and mice (Kojima-Kita et al. 2016; Kuramochi-Miyagawa et al. 2008). Moreover, the PIWI pathway is a relevant

player in imprinting in mice: its components are required for de novo methylation of the ICR of Rasgrf1 (Watanabe et al. 2011) controlling gene expression. Interestingly, piRNAs spanning from a different locus target a transposable element within the non-coding RNA that spans the ICR and is associated with imprinting of the locus. Also, a recent article suggests a role for piRNAs in regulating gene expression (Wu et al. 2018) and establishes a direct role between pachytene piRNAs, spermiogenesis and embryo viability.

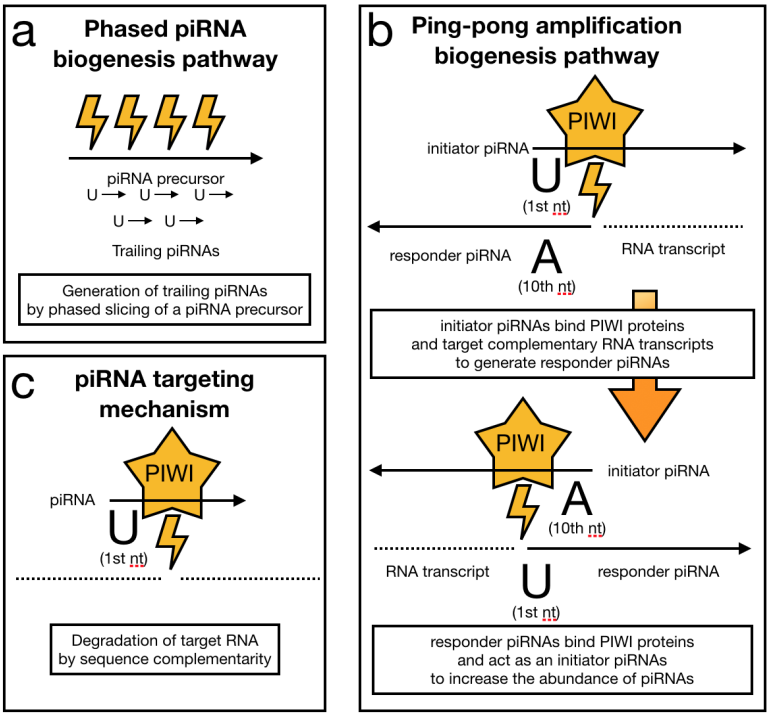


Figure 1. Mechanisms of piRNA generation and transposon targeting. a) Phased piRNAs are generated by an initial trigger and sequentially after that. b) The ping-pong amplification cycle starts with an initiator piRNA targeting an RNA molecule by sequence complementarity and generating piRNAs from it. These responder piRNAs in turn will guide PIWI proteins to other RNA molecules and generate more complementary piRNAs. c) piRNAs sort target RNA molecules for degradation by sequence complementarity.

Transposable elements and the non-coding genome

Transposable elements (TEs) and other repeats make up approximately half of the human and mouse genomes (Mouse Genome Sequencing Consortium 2002) and are the main contributors to different genomic size between species (Sotero-Caio et al. 2017). TEs are mobile elements that contribute to mutagenesis and disease but they are also a source of genomic innovation (reviewed in (Chuong, Elde, and Feschotte 2017). They were first described in maize as controlling elements of genes (Mcclintock 1956).

TE need to be silenced to protect genome integrity (Faulkner et al. 2009). Generally, these regions are compacted in high density chromatin and subjected to a silent state through heterochromatin. Among the mechanisms responsible for the silencing of transposons we find DNA methylation and H3K9me3 (Yoder, Walsh, and Bestor 1997; Walsh, Chaillet, and Bestor 1998; Rowe et al. 2010) in both germline and embryos (reviewed in (Deniz, Frost, and Branco 2019).

Importantly, the genomic context of a gene in terms of potentially regulatory repetitive elements are likely to play a role in its expression (reviewed in (Chuong, Elde, and Feschotte 2017). For instance, TE located 5' of protein-coding loci often act as alternative promoters or tissue-specific transcription start sites (TSS) in a variety of tissues in humans and mice (Faulkner et al. 2009).

Endogenous retroviruses (ERVs), a specific type of transposons, are among the first transcribed sequences during zygotic genome activation in mouse two-cell embryos (Macfarlan et al. 2012). This has a widespread impact on embryos. Specifically, ERVs are essential regulators of development, regulate gene expression networks during embryogenesis in humans (Kunarso et al. 2010) and mice (Macfarlan et al. 2012) and are required in stem cells for totipotency (Fort et al. 2014). ERV are also involved in regulating other complex gene networks of biological processes like dosage-compensation of the X-chromosome in flies (Ellison and Bachtrog 2013) and mammalian immunity (Chuong, Elde, and Feschotte 2016).

Transposon activation can also impact the mammalian transcriptome by triggering DNA methylation of LTR-initiated transcripts in mouse, rat, and human oocytes (Brind'Amour et al. 2018). DNA methylation driven by LTR-initiated transcription, which includes gene promoters, can endure epigenetic reprogramming post fertilization and is associated with transcriptional repression of the maternal allele in adults (Brind'Amour et al. 2018). Hence, variation in LTR insertions can lead to the generation of new imprinted genes.

Some transposable elements like the mouse intracisternal A particle (IAP) family of endogenous retroviruses - the youngest and still active family or ERV-, can transition between active and inactive epigenetic states and are among the few loci that can maintain such state after epigenetic reprogramming (reviewed in Takahashi et al, Cold Spring Harb Symp Quant Biol, 2015). Furthermore, there are examples of IAP-driven heritable expression states (Morgan et al. 1999). For this reason, even though IAPs have been suggested as good candidates for TEI (Lane et al. 2003), a recent study showed most IAPs to reset epigenetic information during embryonic reprogramming (Kazachenka et al. 2018).

Mechanisms of epigenetic inheritance of variable traits through the germline

Authors

Eduard Casas and Tanya Vavouri

Definitions

Epigenetic reprogramming: Refers to the erasure of epigenetic marks such as DNA methylation and histone modifications. During the mammalian life cycle, there are two rounds of reprogramming. In gametogenesis, there is extensive epigenetic reprogramming that includes erasure of genomic imprints. In preimplantation embryonic development there is another round of reprogramming that is linked to the conversion of highly differentiated gametes to the totipotent cells of the early embryo (reviewed in (Reik, Dean, and Walter 2001).

Epialleles: Refers to heritable alternative expression states of a gene that are not caused by genetic variation. Epialleles can often switch states after one or more generations during epigenetic reprogramming or in response to a stimulus (Rakyan et al. 2002).

Paramutation: Process by which one allele of a gene interacts with the other genetically identical allele and affects the latter's expression. Paramutated epialleles are often heritable and can repress other alleles. The mechanism was first discovered in maize (Brink 1956).

Introduction

The idea that non-genetic information can be inherited through the germline was once considered heretical. However, there is a form of inheritance through the germline non dependent on genetic material and generally attributed to epigenetics. In many eukaryotic model organisms, epigenetic information encoded in the germline can transmit variable traits from parental phenotypes to the next generation/s. Inheritance of epigenetic information maintained down to the F2 is considered *transgenerational* inheritance, while we call it *intergenerational* inheritance if the signal is lost in the F1 (Fig 2).

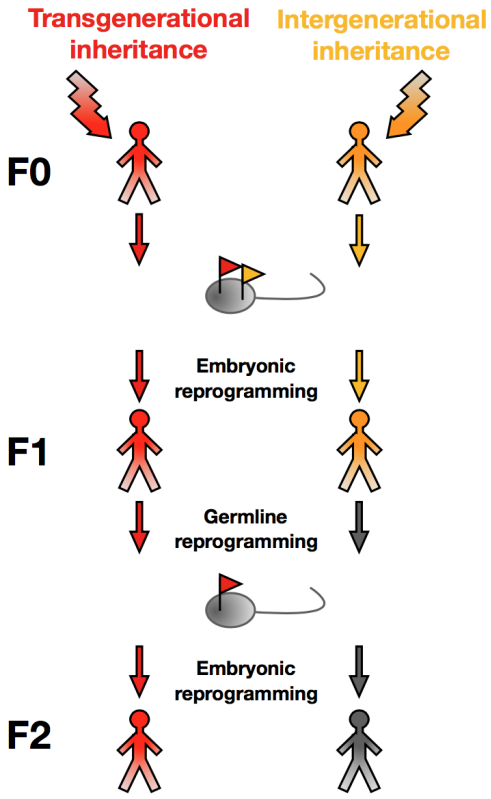


Figure 2. Epigenetic inheritance through the male germline. Intergenerational epigenetic inheritance is inheritance of non-genetic information from parent to offspring through the germline. For information to be inherited between generations it has to escape epigenetic reprogramming in the early embryo. Transgenerational epigenetic inheritance is inheritance of a non-genetic signal from parent to offspring and grand-offspring, without being erased either in the germline or the embryo.

Nowadays, many systems in multiple species suggest different pathways involved in the inheritance of acquired traits. To understand the current state of the field, I review the most striking breakthroughs, some of the best understood cases and the latest research on the epigenetic inheritance field. I discuss the identified mechanisms involved in the transmission of non-genetic information through germline in plants, nematodes, flies and mammals.

A number of mechanisms have been implicated in the transmission of phenotypes between generations but, due to space limitations, in this review we focus only on chromatin, DNA methylation and small non-coding RNAs. Although broadly speaking epigenetic mechanisms are conserved in eukaryotes, there are also important differences. For example, cytosine methylation is widespread in plants and mammals but not in the model organisms *Drosophila melanogaster* (Capuano et

al. 2014) and *Caenorhabditis elegans* (Simpson, Johnson, and Hammen 1986). Also, *C. elegans* spermatozoa, unlike mammalian spermatozoa, do not package their genome in spermatozoa-specific proteins called protamines, potentially allowing the transfer of information between generations through histones (Tabuchi et al. 2018). Therefore, we have organized this review by species. Todate, there are many reported cases of potential epigenetic inheritance in animals (Jablonka and Lamb 1999). For most of these, the mechanism of inheritance is completely unknown. In this review we focus on examples of non-genetic inheritance of variable traits where there is at least some evidence for an underlying epigenetic mechanism, specifically DNA methylation, small RNA or chromatin. We start the review with an overview of epigenetic inheritance in plants, where transgenerational epigenetic inheritance appears to be more common than in animals. We then discuss some of the best-understood models and mechanisms of epigenetic inheritance in nematodes, flies and rodents. Last, we briefly review some notable examples of potential intergenerational responses to environmental exposures in humans.

Models and mechanisms of epigenetic inheritance in plants

Heritable plant epialleles controlled by transposable elements

Vernalization is one of the best-understood examples of how the environment can affect the epigenome of an organism. In *Arabidopsis thaliana*, cold-exposure leads to a progressive deposition of the repressive histone mark H3K27me3 which, after a few weeks, results in silencing of the floral repressor gene (*FLC*). This epigenetic modification is maintained during multiple cell divisions (Gendall et al., 2001; Levy et al., 2002) but is not inherited (Crevillén et al. 2014).

Are there heritable epialleles in plants? Yes, there are multiple (reviewed in (Bond and Baulcombe 2014) and (Quadrana and Colot 2016)). In flowering plants, the germline differentiates from somatic cells which, after being exposed to a particular environmental stimulus (such as cold, drought, etc), can maintain memory of the adaptive response (Gendall et al. 2001; Levy et al. 2002). Also, unlike mammals, plants maintain CG DNA methylation during their life cycle (Calarco et al. 2012). The establishment of DNA methylation involves

small RNAs generated by an RNA dependent RNA polymerase and a Dicer-like protein targeting complementary genomic sites (Bond and Baulcombe 2014). DNA methylation can then be maintained independently of the trigger small RNAs by MET1, the plant homologue of mammalian DNMT1. Therefore in plants, inherited DNA methylation at CG sites provides a simple mechanism to transmit epialleles between generations.

Transposable elements are widespread, can affect gene regulation and maintain memory of expression between generations (Galindo-González et al. 2017). Already in the late 50s, Barbara McClintock, observed that some copies of Suppressor-mutator (*Spm*) transposable elements in maize can reversibly transition between active and inactive states and that these states are heritable (McClintock and Others 1958; McClintock 1961). Transposon activity negatively correlates with DNA methylation (Fedoroff 1989; Martienssen 1998; Banks, Masson, and Fedoroff 1988). Since then, heritable epialleles, often associated with transposable elements, have been identified in many plant and animal species (reviewed in (Slotkin and Martienssen 2007), including species and phenotypes of agricultural and economic value such as the colourless and non-ripening phenotype in tomatoes (Manning et al. 2006), the dwarf phenotype in rice (Miura et al. 2009) and morning glory flowers (Iida et al. 2004).

Association of complex traits with heritable DNA methylation

Is variable DNA methylation the cause of complex, heritable phenotypes? To address this question, a population of inbred *Arabidopsis* plants with highly variable DNA methylation was generated by crossing plants with mutations in two genes involved in DNA methylation (*met1* and *ddm1*) with wild type plants, selecting wild type progeny and then inbreeding (Johannes et al. 2009). Phenotyping and DNA methylation profiling of the resulting epigenetic recombinant inbred lines revealed variation in complex traits such as flowering time and root length associated with the inherited DNA methylation perturbations (Johannes et al. 2009; Cortijo et al. 2014). This experiment provided strong evidence that inherited variation in DNA methylation is causally connected to phenotypic variation in plants.

Models and mechanisms of epigenetic inheritance in *Caenorhabditis elegans*

The selfing mode of reproduction, which reduces genetic diversity, the short generation time and the possibility to keep track of phenotypic changes continuously during many generations makes the nematode worm *Caenorhabditis elegans* an excellent model organism, especially for research on transgenerational epigenetic inheritance. There is no detectable DNA methylation in this animal but small RNA-associated mechanisms and chromatin have been mechanistically linked to transgenerational epigenetic inheritance of gene expression states, including permanent silencing of RNAi-targeted genes (Vastenhouw et al. 2006; Shirayama et al. 2012; Luteijn et al. 2012; Ashe et al. 2012).

Heritable silencing of transgenes, transposons and endogenous loci targeted by small non-coding RNAs

In *C. elegans*, small RNAs play a central role in the establishment, maintenance and transgenerational memory of gene expression (reviewed in (Rechavi and Lev 2017) and (Minkina and Hunter 2018)). Small RNAs that can induce heritable silencing (RNAi) in the nucleus include short interfering RNAs (siRNAs) processed from exogenous double-stranded RNA molecules, endogenous siRNAs and PIWI-interacting RNAs. Argonaute proteins are involved in gene regulation in all species but in *C. elegans* the Argonaute family has massively expanded (Yigit et al. 2006) generating a great diversity of RNA-interacting proteins and types of small regulatory RNAs that are involved in many regulatory mechanisms including epigenetic inheritance. Several members of the Argonaute family and modifiers of repressive (H3K9me3) chromatin have been found to be essential for inter-/trans-generational silencing of transgenes and transposons as well as for the silencing of endogenous genes by exogenous siRNAs. For example, the *C. elegans* PIWI orthologue PRG-1 initiates permanent silencing of transgenes and transposons (Ashe et al. 2012; Shirayama et al. 2012; Lee et al. 2012; Luteijn et al. 2012) and the nuclear Argonaute HRDE-1 is required to transmit memory of gene silencing to the offspring (Buckley et al. 2012; Ashe et al. 2012). Also, conserved H3K9me3 methytransferases are involved in stable silencing of loci targeted by exogenous double-stranded RNA,

transgenes and transposons (Gu et al. 2012; Ashe et al. 2012; Shirayama et al. 2012). Although it is well established that in C. elegans interplay between Argonautes and histone modifiers underly the mechanism of transgenerational silencing triggered by exogenous dsRNA as well as the silencing of transgenes and repeats, it remains debatable whether the molecules that carry the silencing signal in the germline are the small RNAs, the modified histones or both.

Induced chromatin aberrations cause transgenerational phenotypes

In *C. elegans* transgenerational epigenetic inheritance is not limited to RNAi-related repressive mechanisms. Experimentally induced chromatin changes in the germline can cause increasing frequency of sterile offspring, known as the 'mortal germline phenotype' and also transgenerational longevity. For example, animals mutant in either an H3K4 demethylase (Katz et al. 2009) or methyltransferase (Xiao et al. 2011) are initially apparently healthy but give birth to an increasing percentage of sterile offspring after each generation (Katz et al. 2009). Mutations in these chromatin modifiers lead to transgenerational accumulation of chromatin aberrations and increasing gene misregulation during the germline cycle eventually leading to sterility (reviewed in (Kelly 2014)). Interestingly, wild-type offspring of worms mutant in H3K4me3 modifiers have extended longevity for multiple generations (Greer et al. 2011). These studies have thus revealed that errors in chromatin reprogramming and resetting of active chromatin marks during the germline life cycle can have transgenerational consequences not only in the function of the germline but also of somatic cells.

Transgenerational responses to environmental stimuli

In recent years, a number of heritable physiologically acquired epialleles and phenotypes have been discovered in *C. elegans*, including transgenerational responses to diet, temperature, man-made chemicals and inherited behaviours. The mechanisms that have so far been linked to the transmission of these responses through the germline include both chromatin and RNAi.

In *C. elegans*, it has been shown that parental diet affects the phenotype of the future generations. Starvation in young worms leads

to extended lifespan in great-grand-offspring (Rechavi et al. 2014). Inheritance of this phenotype is correlated with differential expression of endogenous siRNAs antisense to a set of protein-coding genes. Inheritance of the differentially expressed endo-siRNAs depends on HRDE-1, a worm specific Argonaute (Rechavi et al. 2014). The endo-siRNAs that are differentially expressed in response to starvation are also bound by HRDE-1 which has led to the hypothesis that endo-siRNAs bound to HRDE-1 carry epigenetic information between generations (Rechavi et al. 2014). On the other end of the nutrition spectrum, worms fed glucose enriched diet appear to have fewer offspring and are resistant to oxidative stress and neurodegeneration with these phenotypes lasting for 1-2 generations (Tauffenberger and Parker 2014). Inheritance of this phenotype requires an intact insulin/IGF-like pathway and components of the H3K4me3 methylation complex (Tauffenberger and Parker 2014). Similarly, exposure to various stressful conditions (arsenite, hyperosmosis and fasting) during development appears to protect offspring from oxidative stress and proteotoxicity (Kishimoto et al. 2017). This inherited response also requires the H3K4me3 complex for transmission of the phenotype to the offspring (Kishimoto et al. 2017). Thus current evidence suggests that inheritance of responses to different diets depends on different epigenetic mechanisms (nuclear RNAi or 'active' histone modifications).

Temperature is another well studied physiological stimulus that leads to transgenerational gene expression changes. Growth at high temperature (25C) for one generation leads to upregulation of endogenous protein-coding genes, transposons and multi-copy transgenes for several generations (Klosin et al. 2017; Schott, Yanai, and Hunter 2014). Endogenous protein-coding genes that maintain a multi-generational memory of expression in response to heat are targeted by endogenous siRNAs (Schott, Yanai, and Hunter 2014). In this model, expression in response to growth in high temperature occurs only through the female germline (Schott, Yanai, and Hunter 2014). Remarkably, when animals are grown at 25C temperature for several generations and then returned to 20C, multi-copy transgene upregulation persists for up to fourteen generations (Klosin et al. 2017). Transgenerational transmission of the active expression state of a multi-copy transgene is inherited through both the male and the

female germline, in *cis* with the locus, is associated with reduced H3K9 methylation at the transgene and requires the worm H3K9 methyltransferase (*set-25*) (Klosin et al. 2017). Similarly, transposons that are normally repressed but upregulated by temperature also maintain memory of temperature-induced activation for several generations (Klosin et al. 2017). The results of these studies suggest that expression changes in response to growth at 25C can be inherited via different mechanisms, depending on the type of gene (transgene, transposon or endogenous protein-coding gene).

It has been reported that exposure to the chemical Bisphenol A (BPA) found in plastics is associated with transgenerational phenotypes. This association was initially observed in rats (Manikkam et al. 2013). In order to gain insight into the possible epigenetic mechanisms responding to BPA, *C. elegans* worms were exposed to BPA and expression of a transgene was measured for several generations (Camacho et al. 2018). This experiment in *C. elegans* revealed that BPA leads to a heritable alteration of heterochromatin that lasts for five generations (Camacho et al. 2018). Inheritance of this response depends on two histone demethylases: the H3K9me3 demethylase JMJ-2 and the H3K27me3 demethylase JMJ-3 (Camacho et al. 2018).

Last but not least, it was recently reported that behavioural responses can also be transgenerationally inherited. Learned avoidance of a bacterial pathogen lasts for four generations (Moore, Kaletsky, and Murphy 2019). Transgenerational avoidance behaviour requires PIWI/PRG-1, MUT-7/RNase D, RRF-1 as well as SET-25 and HPL-2 in the nucleus (Moore, Kaletsky, and Murphy 2019). Especially in the cases of transgenerational changes in behaviour, an important question is how neuronal responses are signalled to the germline. In *C. elegans*, there are two RNA-binding proteins that have so far been associated with neuron-to-germline communication: SID-1 (Devanapally, Ravikumar, and Jose 2015) and RDE-4 (Posner et al. 2019).

Models and mechanisms of epigenetic inheritance in Drosophila melanogaster

Epigenetic inheritance of *Fab-7*-associated epialleles

The polycomb group response element *Fab-7* is one of a few loci known to maintain memory of expression between generations (Cavalli and Paro 1998). The Polycomb group complex binds *Fab-7* and maintains the downstream protein-coding gene in a repressed state during mitotic divisions. Temporary induction of transcription through the *Fab-7* element leads to reporter gene activation that is partially inherited to the offspring (Cavalli and Paro 1998). It was recently shown that a *Fab-7*-controlled transgene, with typically variegated expression, becomes stably active or repressed when one copy of the endogenous *Fab-7* is lost for one generation (Ciabrelli et al. 2017). Transmission of the active or repressed *Fab-7* epiallele for a few generations leads to loss of inter-individual epiallele variation and the emergence of a permanently active or repressed epiallele that is transmitted, unaltered, through the germline (Ciabrelli et al. 2017). Interestingly, stable *Fab-7* epialleles are paramutagenic (Ciabrelli et al. 2017). What is the mechanism of transgenerational transmission of the *Fab-7* epialleles? There is evidence that PRC2 is involved in this phenomenon. Even before the onset of *Fab-7* transgene transcription in the embryo, there is a higher abundance of H3K27me3 at the repressed epiallele than the active epiallele and transient reduction of PRC2 levels leads to faster transition to a stably active epiallele (Ciabrelli et al. 2017). Also, H3K27me3 is inherited through the germline (Zenk et al. 2017). Therefore, PRC2 and the histone modification H3K27me3 are prime candidates for the transmission of acquired epialleles through the germline in flies.

Epigenetic inheritance of environmentally induced responses

Several years ago, a H3K9me3-linked mechanism of epigenetic inheritance was reported in flies. Heat-shock and osmotic stress were shown to lead to disruption of heterochromatin and gene activation that is inherited (Seong et al. 2011). Specifically, the authors reported that stress induces phosphorylation and release of ATF-2 from chromatin, loss of H3K9me2 and activation of a transgene that is

normally repressed by heterochromatin (Seong et al. 2011). The derepressed state is transmitted through both the maternal and the paternal germline, it is paramutagenic and the effect is cumulative (i.e. the number of generations affected is higher if the embryos are under stress for multiple generations) (Seong et al. 2011).

More recently it was also shown that paternal diet leads to metabolic reprogramming in offspring. Specifically, consumption of a high sugar diet by the father leads to increased susceptibility to adiposity in the offspring (Öst et al. 2014). The high amount of sugar in the diet leads to desilencing of a transgene reporter affected by heterochromatin, as well as genes embedded in repressive chromatin (Öst et al. 2014). Interestingly, transcriptional changes detected during embryonic development in the offspring are already forecast in sperm (Öst et al. 2014). Mutant analysis revealed a requirement for an intact heterochromatin machinery both in sperm and in the embryo for transmission of the metabolic phenotype (Öst et al. 2014). Notably, similar transcriptional changes are observed in models of epigenetic obesity in mice and in humans (Öst et al. 2014).

Intergenerational transposon repression by piRNAs

PIWI-interacting RNAs (piRNAs) and piRNA producing loci (piRNA clusters) can also transmit epigenetic information between generations in Drosophila. Maternal piRNAs protect offspring from transposon activation and sterility in crosses of flies with different transposon profiles (Brennecke et al. 2008). Maternally deposited piRNAs targeting transposons in the genome of the offspring are involved in repression of transposon splicing (Teixeira et al. 2017). Although not strictly epigenetic inheritance of an epigenetically variable trait, this demonstrates that maternal small RNAs can affect a phenotype in the offspring. Furthermore, piRNAs can silence homologous loci in *trans* (de Vanssay et al. 2012), akin to paramutation in plants. The trans-silenced homologous loci also produce piRNAs, they become paramutagenic themselves and are heritable across generations (de Vanssay et al. 2012). Furthermore, expression of some piRNAs can be modified by aging and their expression state can be transgenerationally inherited via maternal piRNAs (Grentzinger et al. 2012).

Models and mechanisms of epigenetic inheritance in mice

Intergenerational responses have been studied in rodents for more than forty years. Initially, the focus was on phenotypes transmitted to offspring of female animals exposed to adverse stimuli (e.g. (Cowley and Griesel 1966; Zamenhof, van Marthens, and Grauel 1971; Chandra 1975). There is still scarce evidence that an environmentally induced phenotype can be transmitted for more than two generations in mammals. Nonetheless, by now there is very strong evidence that non-genetic information transmitted through the gametes can influence the phenotype of at least the F1 offspring in mice.

Intergenerational inheritance of IAP-associated epialleles

The most extensively studied model of epigenetic inheritance in mammals is Avy. Avy is a natural allele of the mouse Agouti gene that arose in the C3H/HeJ mouse strain after the insertion of an intracisternal A particle (IAP) repeat upstream of the gene (Dickies 1962). Agouti is a signalling protein that controls melanin production in the hair follicle resulting in agouti coat color. When the Avy IAP is in an active state it acts as a dominant promoter.When the Avy IAP is repressed, the agouti gene is expressed in specific cell types and stages of development from its endogenous promoter and mouse hairs are agouti. When the Avy IAP is active, there is ectopic, constitutive transcription from the LTR of the transposon causing yellow hairs as well as affecting other phenotypes such as obesity, susceptibility to cancer and life expectancy. Thus, the Avy allele exists in two states determined by the state of the IAP. Avy mice can have the Avy allele in either state in all, some or none of their cells. Consequently, genetically identical Avy mice are phenotypically diverse with their colour ranging from yellow through mottled to complete yellow, depending on the activity state of the Avy IAP and the extent of mosaicism.

Avy epialleles are, at least partially, heritable (Wolff et al. 1998; Wolff 1978). Although it was initially thought that transmission of the Avy agouti phenotype from mother to offspring was due to metabolic effects *in utero (Wolff et al. 1998; Wolff 1978)*, oocyte transfer experiments demonstrated that the activity state of the Avy allele

correlates with the activity state in the biological mother and not the phenotype of the foster mother (Morgan et al. 1999). Furthermore, the proportion of Avy mice with agouti colour is higher when both the mother and the grandmother carry the silent Avy epiallele, suggesting accumulation of a signal at the locus (Morgan et al. 1999). Environmental stimuli, in particular maternal diet, can influence the state of the Avy allele in the offspring. Methyl donor supplementation of pregnant mothers (that do not carry the Avy allele themselves) which carry embryos with the Avy allele inherited from the father give birth to more pups with repressed Avy alleles (Cropley et al. 2006). The state of the Avy locus is therefore inherited between generations and responsive to methyl donor supplementation in the diet of the mother.

For a long time, DNA methylation was considered the most likely mechanism of epigenetic inheritance at the Avy locus. The activity state of the Avy epiallele - in both gametes and adult somatic cells - inversely correlates with DNA methylation of the long terminal repeat of the IAP (Morgan et al. 1999; Rakyan et al. 2003; Cooney, Dave, and Wolff 2002; Blewitt et al. 2006). The prevailing hypothesis was that the IAP escapes reprogramming - something that indeed happens at many IAPs (Lane et al. 2003). However, methylation at the Avy IAP, in particular, is erased during mouse preimplantation development (Blewitt et al. 2006). Consequently, the mechanism of maintenance of memory is not simply incomplete erasure of DNA methylation. There is some evidence that polycomb is implicated in intergenerational transmission of Avy epialleles. Although in the C57BL6 mouse genetic background there is no epigenetic inheritance of the Avy epialleles inherited from the father, C57BL6 mice haploinsufficient for Mel18 that inherit the Avy epiallele from an agouti father are more prone to be pseudoagouti than their wild-type littermates (Blewitt et al. 2006). All in all, research on the Avy mouse model has demonstrated strong evidence for intergenerational inheritance through the germline. It is now known that expression of the Avy allele is influenced by the genetic background, the parent from which the allele was inherited, the maternal epigenotype, methyl donor diet supplementation *in utero*, maternal alcohol consumption and even the parents' untransmitted genotypes (Chong et al. 2007; Daxinger and Whitelaw 2012; Kaminen-Ahola et al. 2010; Daxinger et al. 2016). Still, even for

this highly studied locus, the mechanism of transmission of information through the germline remains not fully understood.

Avy is not the only IAP-derived mouse allele with heritable epialleles (see for example (Rakyan et al. 2003). A widespread assumption has been that IAPs have intrinsic features that predispose them to epigenetic metastability. This hypothesis was recently tested using an unbiased screen for IAPs that are variably methylated in individuals of inbred mice, associated with the expression of nearby genes and stably (un)methylated within the same individual (Kazachenka et al. 2018). The results of this screen argue that the activity state of IAPs is not heritable, with very rare exceptions (Kazachenka et al. 2018).

Intergenerational transmission of information via sperm RNA

Currently, a lot of research on epigenetic inheritance in rodents is based on transmission through the male germline. This is because *in utero* effects can be excluded. Additionally, it is easier to limit the interaction of the father with the offspring limiting transmission of phenotypes via altered behaviour. Furthermore, mammalian sperm cells contain very little RNA in comparison to oocytes, making it relatively easier to narrow down the candidate signalling molecules.

The first time it was demonstrated that sperm RNA transferred to a fertilised oocyte can affect an adult phenotype was when sperm RNA from Kittm1Alf/+ mutants microinjected into wild type one-cell embryos was shown to mimic the Kit mutant white-spotted phenotype (Rassoulzadegan et al. 2006). Interestingly, the Kittm1Alf allele, which was engineered by inserting LacZ downstream of the start codon of Kit, can “paramutate” a wild type Kit allele (Rassoulzadegan et al. 2006). Regarding the mechanism of transmission, there is evidence that small RNAs are involved. Microinjection of a variety of RNAs into wild type fertilized oocytes phenocopies the Kit mutation: sperm RNA from mice carrying the paramutated Kit epiallele, microRNAs miR-221, miR-222 and a 28nt oligonucleotide from Kit, all cause embryos to develop white-spots and abnormal Kit transcripts in adulthood (Rassoulzadegan et al. 2006).

Since this initial report, multiple other groups published results showing that microinjections of sperm RNAs into zygotes can induce different phenotypes. For example, sperm small RNAs from affected sires delivered to fertilized eggs can mimic paternal phenotypes including heart hypertrophy (Wagner et al. 2008), general overgrowth (Grandjean et al. 2009), obesity and metabolic disorder (Grandjean et al. 2015; Chen et al. 2016), stress reactivity (Rodgers et al. 2015), enhanced synaptic plasticity (Benito et al. 2018) and traumatic stress (Gapp et al. 2018). Also, sperm microRNAs appear to be important for early embryonic development and implantation (Yuan et al. 2016; Conine et al. 2018). Still, the direct targets of sperm small RNAs in the early embryo and the subsequent cascade of events that leads to the development of a phenotype remain to be discovered.

How does the environment affect sperm RNAs? Profiling of small RNAs from male gametes at different stages of differentiation and maturation from testis and epididymis has revealed that extracellular vesicles can transport small RNAs including tRNA fragments and microRNAs from the epithelium to mature sperm cells (Sharma et al. 2018, 2016). Also high fat diet can induce expression of the RNA methyltransferase Dnmt2 in epididymis, alter RNA modifications and the secondary structure of tRNA fragments (Zhang et al. 2018). Furthermore, DNMT2 is required for epigenetic inheritance and paramutation at the Kit locus (Kiani et al. 2013) and for inheritance of high-fat-diet-induced metabolic disorders (Zhang et al. 2018). Thus, DNMT2 may play a role in encoding external stimuli in the RNA of mature sperm that could influence gene expression in the offspring.

Evidence of epigenetic inheritance in humans

Evidence of epigenetic inheritance in humans is based on epidemiological studies. Whether inheritance of human traits is through epigenetic mechanisms and regulatory signals encoded in germ cells is nearly impossible to prove. Research is focused on identifying associations between the phenotype or the environment of the parents with the phenotype of the offspring. Human sperm cells are easily accessible, therefore it is also relatively easy to identify environmental stimuli significantly affecting the sperm “epigenome” (Casas and Vavouri 2014).

Intergenerational disease susceptibility associated with environmental perturbations

Population-based studies and experiments in mice have repeatedly revealed significant associations between the diet of the parents and metabolic disorders in the offspring (Roseboom and Watson 2012; Jimenez-Chillaron et al. 2009; Pentinat et al. 2010). For instance, children of fathers exposed to famine *in-utero* tend to be heavier and with a higher body-mass index (BMI) than unexposed relatives (Veenendaal et al. 2013). Individuals that were on early stages of gestation during the famine were the most affected. Perhaps this period of embryonic development is a time when environmental perturbations can affect the germline most significantly. Similarly, children who suffered undernutrition *in-utero* during a big famine in China were at a higher risk of developing hyperglycemia and type 2 diabetes as adults (Li et al. 2017). These phenotypic differences were maintained for two generations. Furthermore, a study of three different Överkalix cohorts in Sweden, revealed intergenerational effects of starvation during the first three years of life of grandparents - food shortage correlated with survival in grand-offspring. Strikingly the phenotype was only transmitted to grand-offspring of the same sex, in spite of the signal being transmitted through the same parents (Pembrey et al. 2006).

DNA methylation and small RNA signatures associated with intergenerational effects

Multiple studies have identified associations between somatic DNA methylation variation in offspring and parental exposures. These include methylation variation at the imprinted gene *IGF2* associated with *in utero* undernutrition in the Dutch famine population (Heijmans et al. 2008), methylation variation associated with the season of conception (Waterland et al. 2010; Dominguez-Salas et al. 2014) and methylation variation of the non-coding VTRNA2-1 transcript associated with maternal nutrition (Silver et al. 2015). An increasing body of literature is also revealing associations between the abundance of methylation state of specific loci and small RNAs in sperm and the environment experienced by men. For example, DNA methylation variation and sperm small RNA abundance correlate with

obesity (Donkin et al. 2016; Soubry et al. 2016) and early life stress (Dickson et al. 2018). Furthermore, methylation at genes involved in the control of appetite changes in response to bariatric surgery (Donkin et al. 2016). It remains to be seen whether these changes in DNA methylation and RNA molecules in mature sperm have any direct impact on gene regulation in the human zygote.

Remaining challenges and future perspectives

Carefully performed experiments in model organisms have demonstrated that non-genetic information can be transmitted through the germline and affect the phenotype of the offspring. Epigenetic mechanisms linked to the transmission of information between generations are DNA methylation, RNAi and chromatin. In both plants and animals, a recurring theme is that genomic loci that maintain memory of gene expression state between generations are associated with transposons.

There is currently a huge research effort into better understanding the mechanisms involved in epigenetic inheritance. Important challenges include the dissection of the underlying mechanisms, when perturbation of the candidate mechanisms affects gametogenesis, causing sterility or early embryonic death. Also epigenetic pathway mutant plants or animals often have grossly perturbed epigenomes making it difficult to dissect direct from indirect effects. The development of targeted epigenetic editing of gametes and early embryos, which is starting to become possible (Wei et al. 2019), is going to transform this field in the near future.

Acknowledgements

We thank Adelheid Lempradl and Ben Lehner for comments on the manuscript. We apologize to authors of many relevant papers on the topic that we could not cite due to space limitations. The lab is funded by the Catalan Agency for Management of University and Research Grants [2017 SGR 1262] and the CERCA Programme/Generalitat de Catalunya. Research at the IJC was supported by the “La Caixa” foundation, the Fundación Internacional Josep Carreras and Celgene Spain.

Bibliography

Ashe, Alyson, Alexandra Sapetschnig, Eva-Maria Weick, Jacinth Mitchell, Marloes P. Bagijn, Amy C. Cording, Anna-Lisa Doebley, et al. 2012. “piRNAs Can Trigger a Multigenerational Epigenetic Memory in the Germline of C. Elegans.” *Cell* 150 (1): 88–99.

Banks, J. A., P. Masson, and N. Fedoroff. 1988. “Molecular Mechanisms in the Developmental Regulation of the Maize Suppressor-Mutator Transposable Element.” *Genes & Development* 2 (11): 1364–80.

Benito, Eva, Cemil Kerimoglu, Binu Ramachandran, Tonatiuh Pena-Centeno, Gaurav Jain, Roman Manuel Stilling, Md Rezaul Islam, et al. 2018. “RNA-Dependent Intergenerational Inheritance of Enhanced Synaptic Plasticity after Environmental Enrichment.” *Cell Reports* 23 (2): 546–54.

Blewitt, Marnie E., Nicola K. Vickaryous, Andras Paldi, Haruhiko Koseki, and Emma Whitelaw. 2006. “Dynamic Reprogramming of DNA Methylation at an Epigenetically Sensitive Allele in Mice.” *PLoS Genetics* 2 (4): e49.

Bond, Donna M., and David C. Baulcombe. 2014. “Small RNAs and Heritable Epigenetic Variation in Plants.” *Trends in Cell Biology* 24 (2): 100–107.

Brennecke, Julius, Colin D. Malone, Alexei A. Aravin, Ravi Sachidanandam, Alexander Stark, and Gregory J. Hannon. 2008. “An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing.” *Science* 322 (5906): 1387–92.

Brink, R. A. 1956. “A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible.” *Genetics* 41 (6): 872–89.

Buckley, Bethany A., Kirk B. Burkhart, Sam Guoping Gu, George Spracklin, Aaron Kershner, Heidi Fritz, Judith Kimble, Andrew Fire, and Scott Kennedy. 2012. “A Nuclear Argonaute Promotes Multigenerational Epigenetic Inheritance and Germline Immortality.” *Nature* 489 (7416): 447–51.

Calarco, Joseph P., Filipe Borges, Mark T. A. Donoghue, Frédéric Van Ex, Pauline E. Jullien, Telma Lopes, Rui Gardner, et al. 2012. “Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA.” *Cell* 151 (1): 194–205.

Camacho, Jessica, Lisa Truong, Zeyneb Kurt, Yen-Wei Chen, Marco Morselli, Gerardo Gutierrez, Matteo Pellegrini, Xia Yang, and Patrick Allard. 2018. “The Memory of Environmental Chemical Exposure in C. Elegans Is Dependent on the Jumonji Demethylases Jmjd-2 and Jmjd-3/utx-1.” *Cell Reports* 23 (8): 2392–2404.

Capuano, Floriana, Michael Mülleder, Robert Kok, Henk J. Blom, and Markus Ralser. 2014. “Cytosine DNA Methylation Is Found in Drosophila Melanogaster but Absent in Saccharomyces Cerevisiae, Schizosaccharomyces Pombe, and Other Yeast Species.” *Analytical Chemistry* 86 (8): 3697–3702.

Casas, Eduard, and Tanya Vavouri. 2014. “Sperm Epigenomics: Challenges and Opportunities.” *Frontiers in Genetics* 5 (September): 330.

Cavalli, G., and R. Paro. 1998. “The Drosophila Fab-7 Chromosomal Element Conveys Epigenetic Inheritance during Mitosis and Meiosis.” *Cell* 93 (4): 505–18.

Chandra, R. K. 1975. “Antibody Formation in First and Second Generation Offspring of Nutritionally Deprived Rats.” *Science* 190 (4211): 289–90.

Chen, Qi, Menghong Yan, Zhonghong Cao, Xin Li, Yunfang Zhang, Junchao Shi, Gui-Hai Feng, et al. 2016. “Sperm tsRNAs Contribute to Intergenerational Inheritance of an Acquired Metabolic Disorder.” *Science* 351 (6271): 397–400.

Chong, Suyinn, Nicola Vickaryous, Alyson Ashe, Natasha Zamudio, Neil Youngson, Sarah Hemley, Tomas Stopka, et al. 2007. “Modifiers of Epigenetic Reprogramming Show Paternal Effects in the Mouse.” *Nature Genetics* 39 (5): 614–22.

Ciabrelli, Filippo, Federico Comoglio, Simon Fellous, Boyan Bonev,

Maria Ninova, Quentin Szabo, Anne Xuéreb, et al. 2017. “Stable Polycomb-Dependent Transgenerational Inheritance of Chromatin States in Drosophila.” *Nature Genetics* 49 (6): 876–86.

Conine, Colin C., Fengyun Sun, Lina Song, Jaime A. Rivera-Pérez, and Oliver J. Rando. 2018. “Small RNAs Gained during Epididymal Transit of Sperm Are Essential for Embryonic Development in Mice.” *Developmental Cell* 46 (4): 470–80.e3.

Cooney, Craig A., Apurva A. Dave, and George L. Wolff. 2002. “Maternal Methyl Supplements in Mice Affect Epigenetic Variation and DNA Methylation of Offspring.” *The Journal of Nutrition* 132 (8 Suppl): 2393S – 2400S.

Cortijo, Sandra, René Wardenaar, Maria Colomé-Tatché, Arthur Gilly, Mathilde Etcheverry, Karine Labadie, Erwann Caillieux, et al. 2014. “Mapping the Epigenetic Basis of Complex Traits.” *Science* 343 (6175): 1145–48.

Cowley, J. J., and R. D. Griesel. 1966. “The Effect on Growth and Behaviour of Rehabilitating First and Second Generation Low Protein Rats.” *Animal Behaviour* 14 (4): 506–17.

Crevillén, Pedro, Hongchun Yang, Xia Cui, Christiaan Greeff, Martin Trick, Qi Qiu, Xiaofeng Cao, and Caroline Dean. 2014. “Epigenetic Reprogramming That Prevents Transgenerational Inheritance of the Vernalized State.” *Nature* 515 (7528): 587–90.

Cropley, Jennifer E., Catherine M. Suter, Kenneth B. Beckman, and David I. K. Martin. 2006. “Germ-Line Epigenetic Modification of the Murine Avy Allele by Nutritional Supplementation.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (46): 17308–12.

Daxinger, Lucia, Harald Oey, Luke Isbel, Nadia C. Whitelaw, Neil A. Youngson, Alex Spurling, Kelly K. D. Vonk, and Emma Whitelaw. 2016. “Hypomethylation of ERVs in the Sperm of Mice Haploinsufficient for the Histone Methyltransferase Setdb1 Correlates with a Paternal Effect on Phenotype.” *Scientific Reports* 6 (April): 25004.

Daxinger, Lucia, and Emma Whitelaw. 2012. “Understanding Transgenerational Epigenetic Inheritance via the Gametes in Mammals.” *Nature Reviews. Genetics* 13 (3): 153–62.

Deans, Carrie, and Keith A. Maggert. 2015. “What Do You Mean, ‘Epigenetic’?” *Genetics* 199 (4): 887–96.

Devanapally, Sindhuja, Snusha Ravikumar, and Antony M. Jose. 2015. “Double-Stranded RNA Made in C. Elegans Neurons Can Enter the Germline and Cause Transgenerational Gene Silencing.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (7): 2133–38.

Dickies, Margaret M. 1962. “A NEW VIABLE YELLOW MUTATION IN THE HOUSE MOUSE.” *Journal of Heredity*. https://doi.org/10.1093/oxfordjournals.jhered.a107129.

Dickson, David A., Jessica K. Paulus, Virginia Mensah, Janis Lem, Lorena Saavedra-Rodriguez, Adrienne Gentry, Kelly Pagidas, and Larry A. Feig. 2018. “Reduced Levels of miRNAs 449 and 34 in Sperm of Mice and Men Exposed to Early Life Stress.” *Translational Psychiatry* 8 (1): 101.

Dominguez-Salas, Paula, Sophie E. Moore, Maria S. Baker, Andrew W. Bergen, Sharon E. Cox, Roger A. Dyer, Anthony J. Fulford, et al. 2014. “Maternal Nutrition at Conception Modulates DNA Methylation of Human Metastable Epialleles.” *Nature Communications* 5 (April): 3746.

Donkin, Ida, Soetkin Versteyhe, Lars R. Ingerslev, Kui Qian, Mie Mechta, Loa Nordkap, Brynjulf Mortensen, et al. 2016. “Obesity and Bariatric Surgery Drive Epigenetic Variation of Spermatozoa in Humans.” *Cell Metabolism* 23 (2): 369–78.

Fedoroff, N. V. 1989. “About Maize Transposable Elements and Development.” *Cell* 56 (2): 181–91.

Galindo-González, Leonardo, Corinne Mhiri, Michael K. Deyholos, and Marie-Angèle Grandbastien. 2017. “LTR-Retrotransposons in Plants: Engines of Evolution.” *Gene* 626 (August): 14–25.

Gapp, K., G. van Steenwyk, P. L. Germain, W. Matsushima, K. L. M. Rudolph, F. Manuella, M. Roszkowski, et al. 2018. “Alterations in Sperm Long RNA Contribute to the Epigenetic Inheritance of the Effects of Postnatal Trauma.” *Molecular Psychiatry*, October. https://doi.org/10.1038/s41380-018-0271-6.

Gendall, A. R., Y. Y. Levy, A. Wilson, and C. Dean. 2001. “The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis.” *Cell* 107 (4): 525–35.

Grandjean, Valérie, Sandra Fourré, Diana Andrea Fernandes De Abreu, Marie-Alix Derieppe, Jean-Jacques Remy, and Minoo Rassoulzadegan. 2015. “RNA-Mediated Paternal Heredity of Diet-Induced Obesity and Metabolic Disorders.” *Scientific Reports* 5 (December): 18193.

Grandjean, Valérie, Pierre Gounon, Nicole Wagner, Luc Martin, Kay D. Wagner, Florence Bernex, François Cuzin, and Minoo Rassoulzadegan. 2009. “The miR-124-Sox9 Paramutation: RNA-Mediated Epigenetic Control of Embryonic and Adult Growth.” *Development* 136 (21): 3647–55.

Greer, Eric L., Travis J. Maures, Duygu Ucar, Anna G. Hauswirth, Elena Mancini, Jana P. Lim, Bérénice A. Benayoun, Yang Shi, and Anne Brunet. 2011. “Transgenerational Epigenetic Inheritance of Longevity in Caenorhabditis Elegans.” *Nature* 479 (7373): 365–71.

Grentzinger, Thomas, Claudia Armenise, Christine Brun, Bruno Mugat, Vincent Serrano, Alain Pelisson, and Séverine Chambeyron. 2012. “piRNA-Mediated Transgenerational Inheritance of an Acquired Trait.” *Genome Research* 22 (10): 1877–88.

Gu, Sam Guoping, Julia Pak, Shouhong Guang, Jay M. Maniar, Scott Kennedy, and Andrew Fire. 2012. “Amplification of siRNA in Caenorhabditis Elegans Generates a Transgenerational Sequence-Targeted Histone H3 Lysine 9 Methylation Footprint.” *Nature Genetics* 44 (2): 157–64.

Heijmans, Bastiaan T., Elmar W. Tobi, Aryeh D. Stein, Hein Putter, Gerard J. Blauw, Ezra S. Susser, P. Eline Slagboom, and L. H.

Lumey. 2008. “Persistent Epigenetic Differences Associated with Prenatal Exposure to Famine in Humans.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (44): 17046–49.

Iida, Shigeru, Yasumasa Morita, Jeong-Doo Choi, Kyeung-Il Park, and Atsushi Hoshino. 2004. “Genetics and Epigenetics in Flower Pigmentation Associated with Transposable Elements in Morning Glories.” *Advances in Biophysics* 38: 141–59.

Jablonka, Eva, and Marion J. Lamb. 1999. *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*. Oxford University Press.

Jimenez-Chillaron, Josep C., Elvira Isganaitis, Marika Charalambous, Stephane Gesta, Thais Pentinat-Pelegrin, Ryan R. Faucette, Jessica P. Otis, et al. 2009. “Intergenerational Transmission of Glucose Intolerance and Obesity by in Utero Undernutrition in Mice.” *Diabetes* 58 (2): 460–68.

Johannes, Frank, Emmanuelle Porcher, Felipe K. Teixeira, Vera Saliba-Colombani, Matthieu Simon, Nicolas Agier, Agnès Bulski, et al. 2009. “Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits.” *PLoS Genetics* 5 (6): e1000530.

Kaminen-Ahola, Nina, Arttu Ahola, Murat Maga, Kylie-Ann Mallitt, Paul Fahey, Timothy C. Cox, Emma Whitelaw, and Suyinn Chong. 2010. “Maternal Ethanol Consumption Alters the Epigenotype and the Phenotype of Offspring in a Mouse Model.” *PLoS Genetics* 6 (1): e1000811.

Katz, David J., T. Matthew Edwards, Valerie Reinke, and William G. Kelly. 2009. “A C. Elegans LSD1 Demethylase Contributes to Germline Immortality by Reprogramming Epigenetic Memory.” *Cell* 137 (2): 308–20.

Kazachenka, Anastasiya, Tessa M. Bertozzi, Marcela K. Sjoberg-Herrera, Nic Walker, Joseph Gardner, Richard Gunning, Elena Pahita, Sarah Adams, David Adams, and Anne C. Ferguson-Smith. 2018. “Identification, Characterization, and Heritability of Murine Metastable Epialleles: Implications for Non-Genetic Inheritance.”

Cell 175 (6): 1717.

Kelly, William G. 2014. “Transgenerational Epigenetics in the Germline Cycle of Caenorhabditis Elegans.” *Epigenetics & Chromatin* 7 (1): 6.

Kiani, Jafar, Valérie Grandjean, Reinhard Liebers, Francesca Tuorto, Hossein Ghanbarian, Frank Lyko, François Cuzin, and Minoo Rassoulzadegan. 2013. “RNA-Mediated Epigenetic Heredity Requires the Cytosine Methyltransferase Dnmt2.” *PLoS Genetics* 9 (5): e1003498.

Kishimoto, Saya, Masaharu Uno, Emiko Okabe, Masanori Nono, and Eisuke Nishida. 2017. “Environmental Stresses Induce Transgenerationally Inheritable Survival Advantages via Germline-to-Soma Communication in Caenorhabditis Elegans.” *Nature Communications* 8 (January): 14031.

Klosin, Adam, Eduard Casas, Cristina Hidalgo-Carcedo, Tanya Vavouri, and Ben Lehner. 2017. “Transgenerational Transmission of Environmental Information in C. Elegans.” *Science*. https://doi.org/10.1126/science.aah6412.

Lane, Natasha, Wendy Dean, Sylvia Erhardt, Petra Hajkova, Azim Surani, Jörn Walter, and Wolf Reik. 2003. “Resistance of IAPs to Methylation Reprogramming May Provide a Mechanism for Epigenetic Inheritance in the Mouse.” *Genesis* 35 (2): 88–93.

Lee, Heng-Chi, Weifeng Gu, Masaki Shirayama, Elaine Youngman, Darryl Conte Jr, and Craig C. Mello. 2012. “C. Elegans piRNAs Mediate the Genome-Wide Surveillance of Germline Transcripts.” *Cell* 150 (1): 78–87.

Levy, Yaron Y., Stéphane Mesnage, Joshua S. Mylne, Anthony R. Gendall, and Caroline Dean. 2002. “Multiple Roles of Arabidopsis VRN1 in Vernalization and Flowering Time Control.” *Science* 297 (5579): 243–46.

Li, Jie, Simin Liu, Songtao Li, Rennan Feng, Lixin Na, Xia Chu, Xiaoyan Wu, et al. 2017. “Prenatal Exposure to Famine and the Development of Hyperglycemia and Type 2 Diabetes in

Adulthood across Consecutive Generations: A Population-Based Cohort Study of Families in Suihua, China." *The American Journal of Clinical Nutrition* 105 (1): 221–27.

Luteijn, Maartje J., Petra van Bergeijk, Lucas J. T. Kaaij, Miguel Vasconcelos Almeida, Elke F. Roovers, Eugene Berezikov, and René F. Ketting. 2012. "Extremely Stable Piwi-Induced Gene Silencing in Caenorhabditis Elegans." *The EMBO Journal* 31 (16): 3422–30.

Manikkam, Mohan, Rebecca Tracey, Carlos Guerrero-Bosagna, and Michael K. Skinner. 2013. "Plastics Derived Endocrine Disruptors (BPA, DEHP and DBP) Induce Epigenetic Transgenerational Inheritance of Obesity, Reproductive Disease and Sperm Epimutations." *PloS One* 8 (1): e55387.

Manning, Kenneth, Mahmut Tör, Mervin Poole, Yiguo Hong, Andrew J. Thompson, Graham J. King, James J. Giovannoni, and Graham B. Seymour. 2006. "A Naturally Occurring Epigenetic Mutation in a Gene Encoding an SBP-Box Transcription Factor Inhibits Tomato Fruit Ripening." *Nature Genetics* 38 (8): 948–52.

Martienssen, R. 1998. "Transposons, DNA Methylation and Gene Control." *Trends in Genetics: TIG*.

McClintock, Barbara. 1961. "Some Parallels Between Gene Control Systems in Maize and in Bacteria." *The American Naturalist*. https://doi.org/10.1086/282188.

McClintock, Barbara, and Others. 1958. "The Suppressor-Mutator System of Control of Gene Action in Maize." *Year Book - Carnegie Institution of Washington* 57: 415–29.

Minkina, Olga, and Craig P. Hunter. 2018. "Intergenerational Transmission of Gene Regulatory Information in Caenorhabditis Elegans." *Trends in Genetics: TIG* 34 (1): 54–64.

Miura, Kotaro, Masakazu Agetsuma, Hidemi Kitano, Atsushi Yoshimura, Makoto Matsuoka, Steven E. Jacobsen, and Motoyuki Ashikari. 2009. "A Metastable DWARF1 Epigenetic Mutant Affecting Plant Stature in Rice." *Proceedings of the National*

Academy of Sciences of the United States of America 106 (27): 11218–23.

Moore, Rebecca S., Rachel Kaletsky, and Coleen T. Murphy. 2019. "Piwi/PRG-1 Argonaute and TGF-β Mediate Transgenerational Learned Pathogenic Avoidance." *Cell*, May. https://doi.org/10.1016/j.cell.2019.05.024.

Morgan, H. D., H. G. Sutherland, D. I. Martin, and E. Whitelaw. 1999. "Epigenetic Inheritance at the Agouti Locus in the Mouse." *Nature Genetics* 23 (3): 314–18.

Öst, Anita, Adelheid Lempradl, Eduard Casas, Melanie Weigert, Theodor Tiko, Merdin Deniz, Lorena Pantano, et al. 2014. "Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity." *Cell* 159 (6): 1352–64.

Pembrey, Marcus E., Lars Olov Bygren, Gunnar Kaati, Sören Edvinsson, Kate Northstone, Michael Sjöström, Jean Golding, and ALSPAC Study Team. 2006. "Sex-Specific, Male-Line Transgenerational Responses in Humans." *European Journal of Human Genetics: EJHG* 14 (2): 159–66.

Pentinat, Thais, Marta Ramon-Krauel, Judith Cebria, Ruben Diaz, and Josep C. Jimenez-Chillaron. 2010. "Transgenerational Inheritance of Glucose Intolerance in a Mouse Model of Neonatal Overnutrition." *Endocrinology*. https://doi.org/10.1210/en.2010-0684.

Posner, Rachel, Itai Antoine Toker, Olga Antonova, Ekaterina Star, Sarit Anava, Eran Azmon, Michael Hendricks, Shahar Bracha, Hila Gingold, and Oded Rechavi. 2019. "Neuronal Small RNAs Control Behavior Transgenerationally." *Cell*, June. https://doi.org/10.1016/j.cell.2019.04.029.

Quadrana, Leandro, and Vincent Colot. 2016. "Plant Transgenerational Epigenetics." *Annual Review of Genetics*. https://doi.org/10.1146/annurev-genet-120215-035254.

Rakyan, Vardhman K., Marnie E. Blewitt, Riki Druker, Jost I. Preis, and Emma Whitelaw. 2002. "Metastable Epialleles in Mammals."

Trends in Genetics. https://doi.org/10.1016/s0168-9525(02)02709-9.

Rakyan, Vardhman K., Suyinn Chong, Marnie E. Champ, Peter C. Cuthbert, Hugh D. Morgan, Keith V. K. Luu, and Emma Whitelaw. 2003. "Transgenerational Inheritance of Epigenetic States at the Murine Axin(Fu) Allele Occurs after Maternal and Paternal Transmission." *Proceedings of the National Academy of Sciences of the United States of America* 100 (5): 2538–43.

Rassoulzadegan, Minoo, Valérie Grandjean, Pierre Gounon, Stéphane Vincent, Isabelle Gillot, and François Cuzin. 2006. "RNA-Mediated Non-Mendelian Inheritance of an Epigenetic Change in the Mouse." *Nature* 441 (7092): 469–74.

Rechavi, Oded, Leah Houri-Ze'evi, Sarit Anava, Wee Siong Sho Goh, Sze Yen Kerk, Gregory J. Hannon, and Oliver Hobert. 2014. "Starvation-Induced Transgenerational Inheritance of Small RNAs in C. Elegans." *Cell* 158 (2): 277–87.

Rechavi, Oded, and Itamar Lev. 2017. "Principles of Transgenerational Small RNA Inheritance in Caenorhabditis Elegans." *Current Biology: CB* 27 (14): R720–30.

Reik, W., W. Dean, and J. Walter. 2001. "Epigenetic Reprogramming in Mammalian Development." *Science* 293 (5532): 1089–93.

Rodgers, Ali B., Christopher P. Morgan, N. Adrian Leu, and Tracy L. Bale. 2015. "Transgenerational Epigenetic Programming via Sperm microRNA Recapitulates Effects of Paternal Stress." *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1508347112.

Roseboom, T. J., and E. D. Watson. 2012. "The next Generation of Disease Risk: Are the Effects of Prenatal Nutrition Transmitted across Generations? Evidence from Animal and Human Studies." *Placenta* 33 Suppl 2 (November): e40–44.

Schott, Daniel, Itai Yanai, and Craig P. Hunter. 2014. "Natural RNA Interference Directs a Heritable Response to the Environment." *Scientific Reports* 4 (December): 7387.

Seong, Ki-Hyeon, Dong Li, Hideyuki Shimizu, Ryoichi Nakamura, and Shunsuke Ishii. 2011. “Inheritance of Stress-Induced, ATF-2-Dependent Epigenetic Change.” *Cell* 145 (7): 1049–61.

Sharma, Upasna, Colin C. Conine, Jeremy M. Shea, Ana Boskovic, Alan G. Derr, Xin Y. Bing, Clemence Belleannee, et al. 2016. “Biogenesis and Function of tRNA Fragments during Sperm Maturation and Fertilization in Mammals.” *Science* 351 (6271): 391–96.

Sharma, Upasna, Fengyun Sun, Colin C. Conine, Brian Reichholf, Shweta Kukreja, Veronika A. Herzog, Stefan L. Ameres, and Oliver J. Rando. 2018. “Small RNAs Are Trafficked from the Epididymis to Developing Mammalian Sperm.” *Developmental Cell* 46 (4): 481–94.e6.

Shirayama, Masaki, Meetu Seth, Heng-Chi Lee, Weifeng Gu, Takao Ishidate, Darryl Conte Jr, and Craig C. Mello. 2012. “piRNAs Initiate an Epigenetic Memory of Nonself RNA in the C. Elegans Germline.” *Cell* 150 (1): 65–77.

Silver, Matt J., Noah J. Kessler, Branwen J. Hennig, Paula Dominguez-Salas, Eleonora Laritsky, Maria S. Baker, Cristian Coarfa, et al. 2015. “Independent Genomewide Screens Identify the Tumor Suppressor VTRNA2-1 as a Human Epiallele Responsive to Periconceptional Environment.” *Genome Biology* 16 (June): 118.

Simpson, V. J., T. E. Johnson, and R. F. Hammen. 1986. “Caenorhabditis Elegans DNA Does Not Contain 5-Methylcytosine at Any Time during Development or Aging.” *Nucleic Acids Research* 14 (16): 6711–19.

Slotkin, R. Keith, and Robert Martienssen. 2007. “Transposable Elements and the Epigenetic Regulation of the Genome.” *Nature Reviews. Genetics* 8 (4): 272–85.

Soubry, Adelheid, Lisa Guo, Zhiqing Huang, Cathrine Hoyo, Stephanie Romanus, Thomas Price, and Susan K. Murphy. 2016. “Obesity-Related DNA Methylation at Imprinted Genes in Human Sperm: Results from the TIEGER Study.” *Clinical Epigenetics* 8 (May): 51.

Tabuchi, Tomoko M., Andreas Rechtsteiner, Tess E. Jeffers, Thea A.

Egelhofer, Coleen T. Murphy, and Susan Strome. 2018. “Caenorhabditis Elegans Sperm Carry a Histone-Based Epigenetic Memory of Both Spermatogenesis and Oogenesis.” *Nature Communications* 9 (1): 4310.

Tauffenberger, Arnaud, and J. Alex Parker. 2014. “Heritable Transmission of Stress Resistance by High Dietary Glucose in Caenorhabditis Elegans.” *PLoS Genetics* 10 (5): e1004346.

Teixeira, Felipe Karam, Martyna Okuniewska, Colin D. Malone, Rémi-Xavier Coux, Donald C. Rio, and Ruth Lehmann. 2017. “piRNA-Mediated Regulation of Transposon Alternative Splicing in the Soma and Germ Line.” *Nature* 552 (7684): 268.

Vanssay, Augustin de, Anne-Laure Bougé, Antoine Boivin, Catherine Hermant, Laure Teysset, Valérie Delmarre, Christophe Antoniewski, and Stéphane Ronsseray. 2012. “Paramutation in Drosophila Linked to Emergence of a piRNA-Producing Locus.” *Nature* 490 (7418): 112–15.

Vastenhouw, Nadine L., Karin Brunschwig, Kristy L. Okihara, Fritz Müller, Marcel Tijsterman, and Ronald H. A. Plasterk. 2006. “Gene Expression: Long-Term Gene Silencing by RNAi.” *Nature* 442 (7105): 882.

Veenendaal, M. V. E., R. C. Painter, S. R. de Rooij, P. M. M. Bossuyt, J. A. M. van der Post, P. D. Gluckman, M. A. Hanson, and T. J. Roseboom. 2013. “Transgenerational Effects of Prenatal Exposure to the 1944-45 Dutch Famine.” *BJOG: An International Journal of Obstetrics and Gynaecology* 120 (5): 548–53.

Wagner, Kay D., Nicole Wagner, Hossein Ghanbarian, Valérie Grandjean, Pierre Gounon, François Cuzin, and Minoo Rassoulzadegan. 2008. “RNA Induction and Inheritance of Epigenetic Cardiac Hypertrophy in the Mouse.” *Developmental Cell*. https://doi.org/10.1016/j.devcel.2008.03.009.

Waterland, Robert A., Richard Kellermayer, Eleonora Laritsky, Pura Rayco-Solon, R. Alan Harris, Michael Travisano, Wenjuan Zhang, et al. 2010. “Season of Conception in Rural Gambia Affects DNA

Methylation at Putative Human Metastable Epialleles." *PLoS Genetics* 6 (12): e1001252.

Wei, Yanchang, Jingwen Lang, Qian Zhang, Cai-Rong Yang, Zhen-Ao Zhao, Yixin Zhang, Yanzhi Du, and Yun Sun. 2019. "DNA Methylation Analysis and Editing in Single Mammalian Oocytes." *Proceedings of the National Academy of Sciences of the United States of America*, April. https://doi.org/10.1073/pnas.1817703116.

Wolff, G. L. 1978. "Influence of Maternal Phenotype on Metabolic Differentiation of Agouti Locus Mutants in the Mouse." *Genetics* 88 (3): 529–39.

Wolff, G. L., R. L. Kodell, S. R. Moore, and C. A. Cooney. 1998. "Maternal Epigenetics and Methyl Supplements Affect Agouti Gene Expression in Avy/a Mice." *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 12 (11): 949–57.

Xiao, Yu, Cécile Bedet, Valérie J. P. Robert, Thomas Simonet, Steve Dunkelbarger, Cédric Rakotomalala, Gwen Soete, Hendrik C. Korswagen, Susan Strome, and Francesca Palladino. 2011. "Caenorhabditis Elegans Chromatin-Associated Proteins SET-2 and ASH-2 Are Differentially Required for Histone H3 Lys 4 Methylation in Embryos and Adult Germ Cells." *Proceedings of the National Academy of Sciences of the United States of America* 108 (20): 8305–10.

Yigit, Erbay, Pedro J. Batista, Yanxia Bei, Ka Ming Pang, Chun-Chieh G. Chen, Niraj H. Tolia, Leemor Joshua-Tor, Shohei Mitani, Martin J. Simard, and Craig C. Mello. 2006. "Analysis of the C. Elegans Argonaute Family Reveals That Distinct Argonautes Act Sequentially during RNAi." *Cell* 127 (4): 747–57.

Yuan, Shuiqiao, Andrew Schuster, Chong Tang, Tian Yu, Nicole Ortogero, Jianqiang Bao, Huili Zheng, and Wei Yan. 2016. "Sperm-Borne miRNAs and Endo-siRNAs Are Important for Fertilization and Preimplantation Embryonic Development." *Development* 143 (4): 635–47.

Zamenhof, S., E. van Marthens, and L. Grauel. 1971. “DNA (cell Number) in Neonatal Brain: Second Generation (F2) Alteration by Maternal (F0) Dietary Protein Restriction.” *Science* 172 (3985): 850–51.

Zenk, Fides, Eva Loeser, Rosaria Schiavo, Fabian Kilpert, Ozren Bogdanović, and Nicola Iovino. 2017. “Germ Line-Inherited H3K27me3 Restricts Enhancer Function during Maternal-to-Zygotic Transition.” *Science* 357 (6347): 212–16.

Zhang, Yunfang, Xudong Zhang, Junchao Shi, Francesca Tuorto, Xin Li, Yusheng Liu, Reinhard Liebers, et al. 2018. “Dnmt2 Mediates Intergenerational Transmission of Paternally Acquired Metabolic Disorders through Sperm Small Non-Coding RNAs.” *Nature Cell Biology* 20 (5): 535–40.

Sperm epigenomics: challenges and opportunities

Authors

Eduard Casas and Tanya Vavouri

Abstract

Sperm is a highly differentiated cell type whose function is to deliver a haploid genome to the oocyte. The sperm "epigenomes" were traditionally considered to be insignificant - the sperm is transcriptionally inactive, its genome is packaged in sperm-specific protamine toroids instead of nucleosomes, and its DNA methylation profile is erased immediately post-fertilization. Yet, in recent years there has been an increase in the number of reported cases of apparent epigenetic inheritance through the male germline, suggesting that the sperm epigenome may transmit information between generations. At the same time, technical advances have made the genome-wide profiling of different layers of the sperm epigenome feasible. As a result, a large number of datasets have been recently generated and analyzed with the aim to better understand what non-genetic material is contained within the sperm and whether it has any function post-fertilization. Here, we provide an overview of the current knowledge of the sperm epigenomes as well as the challenges in analysing them and the opportunities in understanding the potential non-genetic carriers of information in sperm.

Introduction

Sperm are highly specialized cells that propagate genetic material from father to offspring. Animal studies suggest that mammalian sperm can transmit non-genetic information across generations. This epigenetic information may alter depending upon the father's environmental exposures. In recent years, the different sperm "epigenomes" have been profiled using high throughput sequencing. Sperm is turning from being one of the most poorly to one of the most intensely profiled cell types (Fig 3).

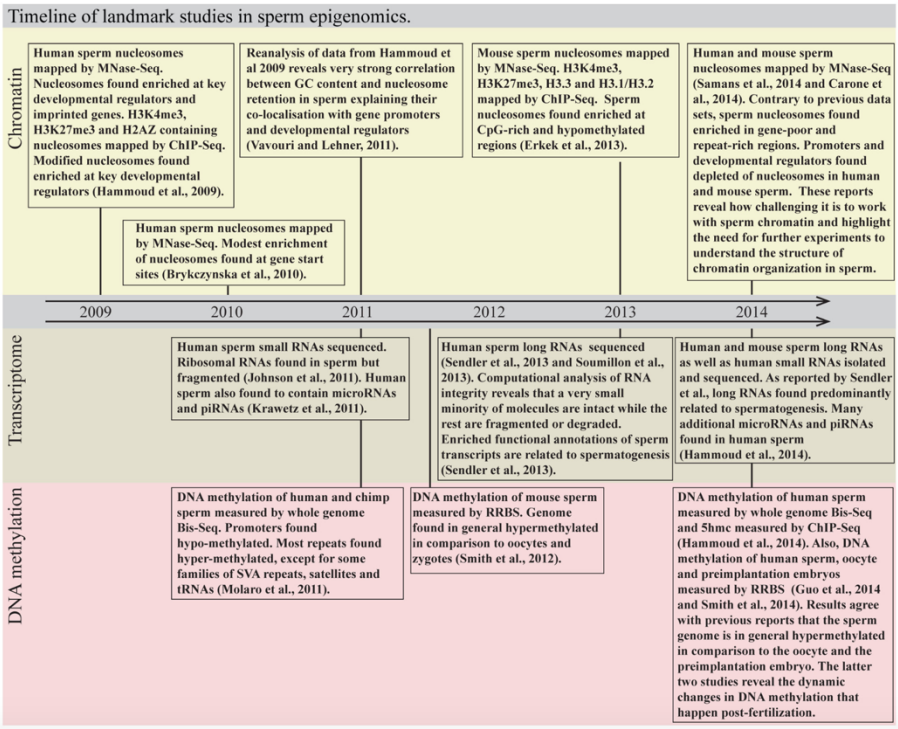


Figure 3. Timeline of landmark studies in sperm epigenomics. MNase-Seq, micrococcal nuclease digestion followed by sequencing; ChIP-Seq, chromatin immunoprecipitation followed by sequencing; Bis-Seq, bisulfite sequencing; RRBS, reduced representation bisulfite sequencing.

Here, we review what is currently known about the RNA, chromatin and DNA methylation profiles of sperm with a focus on human and mouse. We then discuss the experimental and computational challenges in the generation and analysis of sperm epigenome data. Last, we highlight the opportunities raised and the questions that remain unanswered regarding the contents of sperm, especially those related to the impact its non-genetic material has postfertilization.

Sperm transcriptome

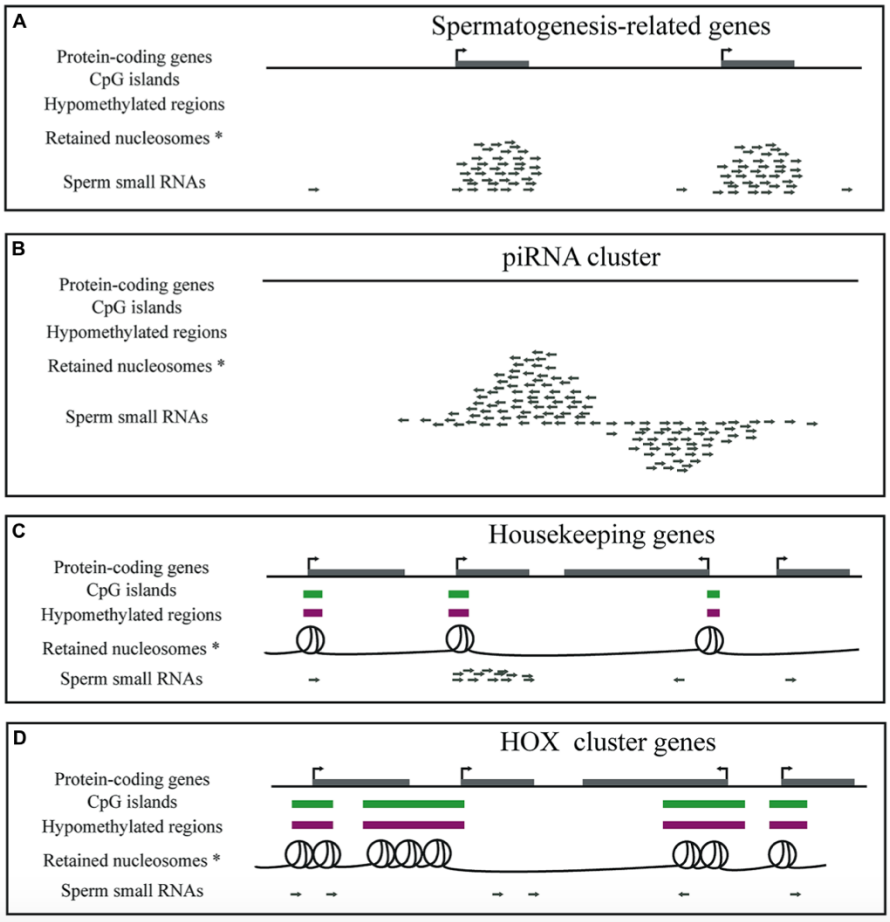


Figure 4. Diagrammatic representation of genes, CpG islands, DNA methylation, nucleosome retention and small RNAs in mature sperm. (A) Sperm cells contain a large number of small RNAs that are fragments of spermatogenesis-related genes, such as the protamine genes. (B) Sperm cells contain piRNAs. (C) GC- and CpG-rich regions overlapping housekeeping gene promoters are hypomethylated and retain nucleosomes in sperm. Small RNA fragments of housekeeping genes expressed until late in sperm development are also present in mature sperm. (D) GC- and CpG-rich regions overlapping developmental regulators, such as the HOX cluster genes, are hypomethylated and retain nucleosomes in sperm. *Note that two of the five genome-wide sperm nucleosome datasets claim that nucleosomes are instead depleted from promoters and enriched at gene poor regions.

Mature sperm cells are transcriptionally inactive (Grunewald et al., 2005; Goodrich et al., 2013). Yet, they do contain RNA (Miller et al., 1994). The vast majority of RNA molecules in sperm are fragments of longer transcripts (Johnson et al., 2011; Sendler et al., 2013; Soumillon et al., 2013; Fig 4A). This includes ribosomal RNA as well as testes and spermatogenesis-specific mRNAs (Johnson et al., 2011). Cessation of transcription and fragmentation of existing sperm mRNAs may be one of the several safety mechanisms that ensure that, upon fertilization, the highly differentiated sperm gives rise to the totipotent zygote. Sperm transcript fragments are an easily accessible record of transcription of the late stages of sperm differentiation and have the potential to be used as markers of fertility (e.g., Yatsenko et al., 2006; Platts et al., 2007).

In addition to fragments of longer transcripts, sperm cells contain a large repertoire of small non-coding RNAs. Like all other cell types, male germ cells express and require the activity of microRNAs (Hayashi et al., 2008; Maatouk et al., 2008; Romero et al., 2011; Wu et al., 2012) and many can still be detected in mature sperm (Amanai et al., 2006; Krawetz et al., 2011; Hammoud et al., 2014). In comparison to oocytes, sperm appears to make an almost insignificant contribution to the total microRNA content of the zygote (Amanai et al., 2006). Nonetheless, at least two different studies have reported that inhibition, in the zygote, of sperm-delivered microRNAs leads to developmental delays (Liu et al., 2012; Hammoud et al., 2014).

Furthermore, dysregulation of at least two different microRNAs (miR-1 and miR-124) in sperm and their transmission to the egg have been postulated to be the causes of two cases of intergenerational inheritance in mouse (Wagner et al., 2008; Grandjean et al., 2009). It should be noted that similar responses were elicited by microinjections of transcript fragments through an unknown mechanism. Also, it was recently shown that traumatic stress in early life of males alters the sperm microRNA (and PIWI-interacting RNA) profile and behavioral and metabolic responses in the offspring (Gapp et al., 2014). These experiments therefore provide evidence that although sperm contains a small quantity of microRNAs in comparison to the oocyte, it still delivers enough to influence preimplantation development and the phenotype of the offspring.

Male germ cells express PIWI-interacting RNAs (piRNAs; Aravin etal., 2006; Girard etal., 2006; Grivna etal., 2006; Lau et al., 2006; Watanabe et al., 2006), also essential small non-coding RNAs for sperm (Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004; Reuter et al., 2011; Fig 4B). A lot remains to be understood about their function, processing and mechanism of action. Their most deeply conserved function is protection of the germline genome from transposons (reviewed in O'Donnell and Boeke, 2007; Thomson and Lin, 2009; Siomi et al., 2011). piRNAs target transposon transcripts for degradation and silencing when DNA methylation (the "default" mechanism of transposon repression) is nearly completely depleted during germ cell development. In addition, a very small number of piRNAs have been linked to imprinting in mouse (Watanabe et al., 2011). Later in sperm development, the role of piRNAs is not as clear, although there is evidence that piRNAs may still protect the genome from transposons (Di Giacomo et al., 2013). Although initially thought to be absent from mature spermatozoa, recent small RNA sequencing studies have revealed more than a thousand known piRNAs from human and mouse sperm samples (Krawetz etal., 2011; Hammoud et al., 2014). The role, if any, of piRNAs in mature sperm is currently unknown. It is also not known whether mature sperm piRNAs are intact and still bound to functional PIWI proteins and whether they have any role in transcriptional or post-transcriptional regulation in the early embryo.

Mature sperm cells contain a plethora of other small RNAs that we currently know little about. There are tRNA fragments that are 30–34 nt long, i.e., the size of piRNAs (Peng et al., 2012), small RNAs processed from piRNA clusters that are 20–21 nt long, (instead of the expected ~30 nt of piRNAs in late spermatogenesis; Kawano etal., 2012) and fragments of repeats (Krawetz et al., 2011). Short transcripts derived from LINE-1 elements were recently found to positively regulate expression of LINE-1 repeats in early mouse embryos (Fadloun et al., 2013), so it is possible that among these fragments there are functional regulatory RNAs. Last, RNA molecules themselves (e.g., tRNAs) can carry modifications (Torres et al., 2014) that have been postulated to carry epigenetic information from father to offspring (Kiani et al., 2013).

Sperm chromatin

Sperm chromatin is highly specialized and is the end product of a highly complex differentiation program during which an impressive number of different testis-specific histone variants, histone-to-protamine transition proteins, and protamine genes are expressed. The role of many of these histone variants and histone-like proteins on gene expression during sperm differentiation and on mature sperm chromatin organization remains to be worked out. Post-fertilization, protamines are released from the paternal genome and replaced by maternal histones [for extensive reviews on protamines see (Lewis et al., 2003; Oliva, 2006; Balhorn, 2007; Rathke et al., 2014)].

In humans, 4–15% of the genome retains histones in sperm (Gatewood et al., 1987; Hammoud et al., 2009). Since the late eighties, it has been known that sperm nucleosomes are not randomly distributed along the genome (Gatewood etal., 1987). Comparing chromatin organization at the globin and protamine genes in sperm samples from different individuals, Gardiner-Garden and colleagues found that it is conserved between individuals (Gardiner-Garden et al., 1998). Interestingly, they also noted that some genes expressed early in development are packaged in nucleosomes while others expressed later are packaged in protamine toroids. Based on these, it was proposed that nucleosomes retained in sperm likely have a structural or regulatory role in late spermiogenesis and/or early embryo development.

Since 2009, several genome-wide sperm nucleosome profiles have been generated (Arpanahi et al., 2009; Hammoud et al., 2009; Brykczynska et al., 2010; Carone et al., 2014; Erkek et al., 2014; Samans et al., 2014). These confirmed that indeed the sites that remain packaged in nucleosomes are not randomly distributed along the genome. The first two studies (Arpanahi et al., 2009; Hammoud et al., 2009) showed that sperm nucleosomes are highly enriched at regulatory regions and in particular overrepresented at genes that regulate embryonic development such as the HOX genes (Hammoud et al., 2009; Fig 4C, D). This result is in agreement with the pre-existing notion that histones in sperm facilitate transcription regulation in the early embryo (Gatewood et al., 1987).

The availability of genome-wide profiles of histone enriched DNA in sperm made it possible to begin to dissect the mechanisms that determine which sites remain packaged by histones and which ones are replaced by protamines (Vavouri and Lehner, 2011; Erkek et al., 2014). Promoters of housekeeping genes and developmental regulators were found to overlap CpG islands, regions with high GC and CpG-content (reviewed in Deaton and Bird, 2011). Indeed, on a genome-wide scale and considering the non-repetitive parts of the genome, that pose problems when dealing with mapping sequenced reads, GC-content showed very strong correlation with histone retention in sperm (Vavouri and Lehner, 2011). This would suggest that the mechanism of nucleosome retention in sperm is tightly associated with sequence composition. Importantly, GC-content was more recently also confirmed to be strongly associated with histone retention in mouse sperm (Erkek et al., 2014). Considering all possible dinucleotides, Erkek etal. (2014) further found that, in mouse, it is CpG-dinucleotide composition that correlates best with the sites that retain histones in sperm. Also, according to both Hammoud et al. (2009) and Erkek et al. (2014), sites that retain histones in sperm are in general hypomethylated, however, it is unclear at this point whether this is due to a direct mechanistic link between DNA methylation and histone retention or whether they simply co-occur at CpG-rich regions.

Surprisingly, the two datasets published in 2014 show very different nucleosome distribution in human and mouse sperm (Carone et al., 2014; Samans et al., 2014). They show nucleosomes preferentially enriched at gene-poor/repeat-rich regions of the genome. Clearly, the six currently available genome-wide datasets of human and mouse sperm nucleosomes cannot all reflect the chromatin structure of sperm. Most likely, there is a critical step in sperm chromatin preparation and even slight variations in the protocol lead to isolation of very different fractions of the genome. According to Carone et al. (2014), this crucial step is the concentration of micrococcal nuclease. However, Samans et al. (2014) apparently used the protocol of Hammoud et al. (2009) but got the opposite results. A systematic comparison of the different sperm nucleosome isolation protocols and comparative analysis of the resulting data remains to be done to

convincingly show what is really the organization of retained nucleosomes in mature sperm.

Sperm histones, like somatic histones, carry posttranslational modifications. Of particular interest, due to their important role in normal development and link with the maintenance of transcription patterns are the trithorax mark histone H3 lysine 4 trimethylation (H3K4me3) and the polycomb mark histone H3 lysine 27 trimethylation (H3K27me3). Sperm chromatin contains both of these (Hammoud etal., 2009; Brykczynska et al., 2010; Erkek et al., 2014). H3K4me3 is enriched at promoters of highly expressed genes during spermatogenesis (Hammoud et al., 2009; Brykczynska et al., 2010). It has also been reported that H3K4me3 marks some of the HOX cluster genes and paternally expressed imprinted genes (Hammoud etal., 2009). H3K27me3 marks primarily developmental regulators such as the HOX genes (Hammoud et al., 2009).

The genome-wide profiles of two histone variants are currently available for sperm. The histone variant H2AZ, which is associated with active regulatory regions in somatic cells, is limited to pericentric heterochromatin in mature sperm (Hammoud et al., 2009). H2AZ is, however, present at promoters of expressed genes in round spermatids (Soboleva et al., 2012; Hammoud et al., 2014). Since (according to the data from Hammoud et al., 2009 and Erkek et al., 2014) many promoters retain nucleosomes in sperm, it is unclear whether H2AZ-containing nucleosomes are lost from promoters in elongating spermatids before the histone-to-protamine transition, or whether nucleosomes lacking this histone variant replace existing nucleosomes during the histone-to-protamine transition. Unlike H2AZ, the histone variant H3.3 is found at expressed genes in round spermatids and is retained at the same promoters in mature sperm (Erkek et al., 2014).

Paternal histones can still be found in the zygote several hours post-fertilization in both human and mouse (van der Heijden et al., 2006, 2008; Puschendorf et al., 2008). To what extend and how exactly paternal histones contribute to chromatin organization and gene expression in the early embryo is not yet clear.

Sperm DNA methylation

Most of the genome of mature sperm is highly methylated (Molaro et al., 2011). This is in stark contrast to the globally lowly methylated oocytes and early embryos (Smallwood et al., 2011; Smith et al., 2012). However, CpG islands including those overlapping developmental regulators such as the HOX genes are hypomethylated (Hammoud et al., 2009; Fig 4). In contrast, promoters of key pluripotency regulators such as those of Oct4 and Nanog are highly methylated in human sperm (Hammoud et al., 2009). In light of these results, the relationship between DNA methylation in sperm and timing of expression in the early embryo is unclear.

The male germline goes through two waves of nearly complete DNA methylation erasure. One of these happens in the zygote, shortly after fertilization. At this stage, DNA methylation is erased specifically from the paternal genome (Oswald et al., 2000; Smith et al., 2012), affecting the majority of the genome but sparing paternal imprints (Edwards and Ferguson-Smith, 2007; Hajkova, 2011; Smith et al., 2012; Hackett and Surani, 2013). This ensures that DNA methylation gained by germ cells during the lifetime of the father is removed before the embryo starts development (Hajkova et al., 2002).

A small number of highly methylated regions, mostly associated with repeats, do nonetheless escape DNA methylation reprogramming. The most prominent example is the mouse IAP family of repeats (Howlett and Reik, 1991; Morgan et al., 1999; Lane et al., 2003; Kim et al., 2004; Guibert et al., 2012; Seisenberger etal., 2012). The mechanism that allows IAPs to evade DNA demethylation is currently unknown.

In search for molecular carriers of non-genetic information from father to offspring, DNA methylation analyses of sperm cells have featured prominently. At least in one study, changes in DNA methylation in sperm did indeed correlate with inheritance of a phenotype (Martinez et al., 2014), although the DNA methylation variation detected in sperm from different fathers was small and could be downstream of the cause of transmission of the phenotype. Not surprisingly, the strongest evidence of DNA methylation variation in

sperm influencing phenotypic variation in offspring is related to IAP elements in mice (Morgan et al., 1999; Rakyan et al., 2003; Blewitt et al., 2006).

Challenges and opportunities lying ahead

Analyzing the sperm transcriptome poses several experimental and computational challenges. The first challenge is that sperm cells have very little RNA. It has been estimated that there are only 10–100 fg of total RNA per human sperm cell (Pessot etal., 1989; Krawetz, 2005), which is much less than that in somatic cells. Consequently, contamination of a sperm sample by somatic cells can heavily bias the resulting RNA profile. The second challenge is the absence of intact ribosomal RNA (Johnson et al., 2011; Goodrich et al., 2013). Quality metrics based on the "intactness" of ribosomal RNA (used for somatic samples) do not apply although they could be used to assess somatic cell contamination. The third challenge is at the analysis stage. Transcript abundance quantification assumes that transcripts are intact. However, in sperm samples, only a tiny fraction (if any) of sequenced reads mapping to a gene correspond to intact transcripts.

The mechanisms and dynamics of sperm transcript fragmentation/degradation are unknown. Until we have a better understanding of these processes and a systematic assessment of how accurately different gene expression quantification methods perform on sperm samples, we need to be cautious interpreting apparent abundance differences between different genes in the same sample and between samples. Transcript fragments also complicate the analysis of small RNAs. Degradation intermediates of ribosomal, mRNA and other transcripts largely outnumber sequence reads mapping to microRNAs. Although these reads can easily be identified and excluded, they also consume a very large proportion of the sequenced reads. So, if somatic small RNA samples can be profiled with as few as 5 million reads, sperm samples require several fold higher numbers of reads to achieve comparable depth of known regulatory small non-coding RNAs.

Analyzing sperm chromatin also poses great challenges. Because it is extremely compacted by protamines instead of histones one needs to

use modified micrococcal nuclease digestion or chromatin immunoprecipitation protocols (e.g., Hammoud et al., 2009; Hisano et al., 2013; Carone et al., 2014). And because the different experimental protocols for protamine-compacted genomes have been less extensively used than those for histone-compacted genomes, their biases are also less understood. For example, as mentioned above, the recent genome-wide profiles of human and mouse sperm nucleosomes arrived to contradicting conclusions (Hammoud etal., 2009; Brykczynska etal., 2010; Hisano etal., 2013; Carone et al., 2014; Samans et al., 2014).

The most fundamental question regarding the transcriptome, chromatin and DNA methylation of sperm is whether they can transmit information about the father's environmental exposures to the offspring. There are currently many reported cases of epigenetic inheritance via sperm (reviewed in Rando, 2012). For example, the father's diet and traumatic experiences in early life seem to influence the phenotype of the offspring (Anderson et al., 2006; Carone et al., 2010; Ng et al., 2010; Gapp et al., 2014; Martinez et al., 2014). Although in some cases candidate carriers of this information have been identified (e.g., RNA or DNA methylation), the mechanisms are far from being adequately understood. Until mechanisms of epigenetic inheritance from father to offspring have been worked out and genetic inheritance has been definitively ruled out, it will remain questionable whether transand inter-generational epigenetic inheritance of phenotypes indeed exists in mammals (Heard and Martienssen, 2014).

Conclusion

Although small, transcriptionally inert, with extremely compacted genome and virtually no cytoplasm, the sperm cell contains a plethora of small RNAs, a large number of DNA sequences packaged by histones and a distinctive DNA methylation profile. Until recently, the main purpose for studying the RNA, chromatin and DNA methylation of sperm (other than scientific curiosity for this highly peculiar cell type) was to identify potential biomarkers of male infertility. Today, there is an additional focus and this is to understand whether any of these "epigenomes" can transmit information from father to offspring.

Therefore, it is now even more important to understand what information these epigenomes contain, how they are set, how they vary between individuals as well as between individual sperm cells, whether they are delivered to the egg upon fertilization and whether they have any impact on the development of the embryo and the phenotype of the offspring. During the past 5 years impressive advances have been made in describing the non-genetic contents of human sperm. Great opportunities are now lying ahead to also understand the mechanisms that set them and whether (and how) they influence gene and genome regulation in the early embryo.

Acknowledgments

The work in our lab is supported by Spanish Ministry grant BFU2011-30246, Ramon y Cajal grant RYC-2010-07114, Marie Curie European Reintegration Grant "Evo-Chromo" and by the IMPPC. Our lab is a member of the European Network of Excellence "EpiGeneSys".

References

Amanai, M., Brahmajosyula, M., and Perry, A. C. (2006). A restricted role for sperm-borne microRNAs in mammalian fertilization. Biol. Reprod. 75, 877–884. doi: 10.1095/biolreprod.106.056499

Anderson, L. M., Riffle, L., Wilson, R., Travlos, G. S., Lubomirski, M. S., and Alvord, W. G. (2006). Preconceptional fasting of fathers alters serum glucose in offspring of mice. Nutrition 22, 327–331. doi: 10.1016/j.nut.2005.09.006

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442, 203–207. doi: 10.1038/nature04916

Arpanahi, A., Brinkworth, M., Iles, D., Krawetz, S. A., Paradowska, A., Platts, A. E., et al. (2009). Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. Genome Res. 19, 1338–1349. doi: 10.1101/gr.094953.109

Balhorn, R. (2007). The protamine family of sperm nuclear proteins. Genome Biol. 8, 227. doi: 10.1186/gb-2007-8-9-227

Blewitt, M. E., Vickaryous, N. K., Paldi, A., Koseki, H., and Whitelaw, E. (2006). Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. PLoS Genet. 2:e49. doi: 10.1371/journal.pgen.0020049

Brykczynska, U., Hisano, M., Erkek, S., Ramos, L., Oakeley, E. J., Roloff, T. C., et al. (2010). Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat. Struct. Mol. Biol. 17, 679–687. doi: 10.1038/nsmb.1821

Carone, B. R., Fauquier, L., Habib, N., Shea, J. M., Hart, C. E., Li, R., et al. (2010). Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. Cell 143, 1084–1096. doi: 10.1016/j.cell.2010.12.008

Carone, B. R., Hung, J. H., Hainer, S. J., Chou, M. T., Carone, D. M., Weng, Z., et al. (2014). High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. Dev. Cell 30, 11–22. doi: 10.1016/j.devcel.2014.05.024

Deaton, A. M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. 25, 1010–1022. doi: 10.1101/gad.2037511

Deng, W., and Lin, H. (2002). Miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev. Cell 2, 819–830. doi: 10.1016/S1534-5807(02)00165-X

Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., et al. (2013). Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol. Cell 50, 601–608. doi: 10.1016/j.molcel.2013.04.026

Edwards, C. A., and Ferguson-Smith, A. C. (2007). Mechanisms regulating imprinted genes in clusters. Curr. Opin. Cell Biol. 19, 281–289. doi: 10.1016/j.ceb.2007.04.013

Erkek, S., Hisano, M., Liang, C.-Y., Gill, M., Murr, R., Dieker, J., etal. (2014). Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. Nat. Struct. Mol. Biol. 20, 868–875. doi: 10.1038/nsmb.2599

Fadloun, A., Le Gras, S., Jost, B., Ziegler-Birling, C., Takahashi, H., Gorab, E., et al. (2013). Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. Nat. Struct. Mol. Biol. 20, 332–338. doi: 10.1038/nsmb.2495

Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, P., Prados, J., et al. (2014). Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. Nat. Neurosci. 17, 667–669. doi: 10.1038/nn.3695

Gardiner-Garden, M., Ballesteros, M., Gordon, M., and Tam, P. P. (1998). Histoneand protamine-DNA association: conservation of different patterns within the beta-globin domain in human sperm. Mol. Cell. Biol. 18, 3350–3356.

Gatewood, J. M., Cook, G. R., Balhorn, R., Bradbury, E. M., and Schmid, C. W. (1987). Sequence-specific packaging of DNA in human sperm chromatin. Science 236, 962–964. doi: 10.1126/science.3576213

Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 442, 199–202. doi: 10.1038/nature04917

Goodrich, R. J., Anton, E., and Krawetz, S. A. (2013). Isolating mRNA and small noncoding RNAs from human sperm. Methods Mol. Biol. 927, 385–396. doi: 10.1007/978-1-62703-038-0_33

Grandjean, V., Gounon, P., Wagner, N., Martin, L., Wagner, K. D., Bernex, F., etal. (2009). The miR-124-Sox9 paramutation: RNA-mediated epigenetic control of embryonic and adult growth. Development 136, 3647–3655. doi: 10.1242/dev.041061

Grivna, S. T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. 20, 1709–1714. doi: 10.1101/gad.1434406

Grunewald, S., Paasch, U., Glander, H. J., and Anderegg, U. (2005). Mature human spermatozoa do not transcribe novel RNA. Andrologia 37, 69–71. doi: 10.1111/j.1439-0272.2005.00656.x

Guibert, S., Forne, T., and Weber, M. (2012). Global profiling of DNA methylation erasure in mouse primordial germ cells. Genome Res. 22, 633–641. doi: 10.1101/gr.130997.111

Hackett, J. A., and Surani, M. A. (2013). DNA methylation dynamics during the mammalian life cycle. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368, 20110328. doi: 10.1098/rstb.2011.0328

Hajkova, P. (2011). Epigenetic reprogramming in the germline: towards the ground state of the epigenome. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366, 2266–2273. doi: 10.1098/rstb.2011.0042

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., et al. (2002).Epigenetic reprogramming in mouse primordial germ cells. Mech. Dev. 117, 15–23. doi: 10.1016/S0925-4773(02)00181-8

Hammoud, S. S., Low, D. H., Yi, C., Carrell, D. T., Guccione, E., and Cairns, B. R. (2014). Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. Cell Stem Cell 15, 239–253. doi: 10.1016/j.stem.2014.04.006

Hammoud, S. S., Nix, D. A., Zhang, H., Purwar, J., Carrell, D. T., and Cairns, B. R. (2009). Distinctive chromatin in human sperm packages genes for embryo development. Nature 460, 473–478. doi: 10.1038/nature08162

Hayashi, K., Chuva De Sousa Lopes, S. M., Kaneda, M., Tang, F., Hajkova, P., Lao, K., etal. (2008). MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. PLoS ONE 3:e1738. doi: 10.1371/journal.pone.0001738

Heard, E., and Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157, 95–109. doi: 10.1016/j.cell.2014.02.045

Hisano, M., Erkek, S., Dessus-Babus, S., Ramos, L., Stadler, M. B., and Peters, A. H. (2013). Genome-wide chromatin analysis in mature mouse and human spermatozoa. Nat. Protoc. 8, 2449–2470. doi: 10.1038/nprot.2013.145

Howlett, S. K., and Reik, W. (1991). Methylation levels of maternal and paternal genomes during preimplantation development. Development 113, 119–127.

Johnson, G. D., Sendler, E., Lalancette, C., Hauser, R., Diamond, M. P., and Krawetz, S. A. (2011). Cleavage of rRNA ensures translational cessation in sperm at fertilization. Mol. Hum. Reprod. 17, 721–726. doi: 10.1093/molehr/gar054

Kawano, M., Kawaji, H., Grandjean, V., Kiani, J., and Rassoulzadegan, M. (2012). Novel small noncoding RNAs in mouse spermatozoa, zygotes and early embryos. PLoS ONE 7:e44542. doi: 10.1371/journal.pone.0044542

Kiani, J., Grandjean, V., Liebers, R., Tuorto, F., Ghanbarian, H., Lyko, F., etal. (2013). RNA-mediated epigenetic heredity requires the cytosine methyltransferase Dnmt2. PLoS Genet. 9:e1003498. doi: 10.1371/journal.pgen.1003498

Kim, S. H., Kang, Y. K., Koo, D. B., Kang, M. J., Moon, S. J., Lee, K. K., et al. (2004). Differential DNA methylation reprogramming of various repetitive sequences in mouse preimplantation embryos. Biochem. Biophys. Res. Commun. 324, 58–63. doi: 10.1016/j.bbrc.2004.09.023

Krawetz, S. A. (2005). Paternal contribution: new insights and future challenges. Nat. Rev. Genet. 6, 633–642. doi: 10.1038/nrg1654

Krawetz, S. A., Kruger, A., Lalancette, C., Tagett, R., Anton, E., Draghici, S., et al. (2011). A survey of small RNAs in human sperm. Hum. Reprod. 26, 3401–3412. doi: 10.1093/humrep/der329

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T. W., Isobe, T., Asada, N., Fujita, Y., et al. (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development 131, 839–849. doi: 10.1242/dev.00973

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J., et al. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 35, 88–93. doi: 10.1002/gene.10168

Lau, N. C., Seto, A. G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D. P., et al. (2006). Characterization of the piRNA complex from rat testes. Science 313, 363–367. doi: 10.1126/science.1130164

Lewis, J. D., Song, Y., De Jong, M. E., Bagha, S. M., and Ausio, J. (2003). A walk though vertebrate and invertebrate protamines. Chromosoma 111, 473–482. doi: 10.1007/s00412-002-0226-0

Liu, W. M., Pang, R. T., Chiu, P. C., Wong, B. P., Lao, K., Lee, K. F., et al. (2012). Sperm-borne microRNA-34c is required for the first cleavage division in mouse. Proc. Natl. Acad. Sci. U.S.A. 109, 490–494. doi: 10.1073/pnas.1110368109

Maatouk, D. M., Loveland, K. L., McManus, M. T., Moore, K., and Harfe, B. D. (2008). Dicer1 is required for differentiation of the mouse male germline. Biol. Reprod. 79, 696–703. doi: 10.1095/biolreprod.108.067827

Martinez, D., Pentinat, T., Ribo, S., Daviaud, C., Bloks, V. W., Cebria, J., et al. (2014). In utero undernutrition in male mice programs liver lipid metabolism in the second-generation offspring involving altered lxra DNA methylation. Cell Metab. 19, 941–951. doi: 10.1016/j.cmet.2014.03.026

Miller, D., Tang, P. Z., Skinner, C., and Lilford, R. (1994). Differential RNA fingerprinting as a tool in the analysis of spermatozoal gene expression. Hum. Reprod. 9, 864–869.

Molaro, A., Hodges, E., Fang, F., Song, Q., McCombie, W. R., Hannon, G. J., et al. (2011). Sperm methylation profiles reveal features of

epigenetic inheritance and evolution in primates. Cell 146, 1029–1041. doi: 10.1016/j.cell.2011.08.016

Morgan, H. D., Sutherland, H. G., Martin, D. I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. Nat. Genet. 23, 314–318. doi: 10.1038/15490

Ng, S. F., Lin, R. C., Laybutt, D. R., Barres, R., Owens, J. A., and Morris, M. J. (2010). Chronic high-fat diet in fathers programs beta-cell dysfunction in female rat offspring. Nature 467, 963–966. doi: 10.1038/nature09491

O'Donnell, K. A., and Boeke, J. D. (2007). Mighty piwis defend the germline against genome intruders. Cell 129, 37–44. doi: 10.1016/j.cell.2007.03.028

Oliva, R. (2006). Protamines and male infertility. Hum. Reprod. Update 12, 417–435. doi: 10.1093/humupd/dml009

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., et al. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. 10, 475–478. doi: 10.1016/S0960-9822(00)00448-6

Peng, H., Shi, J., Zhang, Y., Zhang, H., Liao, S., Li, W., et al. (2012). A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. Cell Res. 22, 1609–1612. doi: 10.1038/cr.2012.141

Pessot, C. A., Brito, M., Figueroa, J., Concha, I. I., Yanez, A., and Burzio, L. O. (1989). Presence of RNA in the sperm nucleus. Biochem. Biophys. Res. Commun. 158, 272–278. doi: 10.1016/S0006-291X(89)80208-6

Platts, A. E., Dix, D. J., Chemes, H. E., Thompson, K. E., Goodrich, R., Rockett, J. C., etal. (2007). Success and failure in human spermatogenesis as revealed by teratozoospermic RNAs. Hum. Mol. Genet. 16, 763–773. doi: 10.1093/hmg/ddm012

Puschendorf, M., Terranova, R., Boutsma, E., Mao, X., Isono, K., Brykczynska, U., etal. (2008). PRC1 and Suv39h specify parental

asymmetry at constitutive heterochromatin in early mouse embryos. Nat. Genet. 40, 411–420. doi: 10.1038/ng.99

Rakyan, V. K., Chong, S., Champ, M. E., Cuthbert, P. C., Morgan, H. D., Luu, K. V., et al. (2003). Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. Proc. Natl. Acad. Sci. U.S.A. 100, 2538–2543. doi: 10.1073/pnas.0436776100

Rando, O. J. (2012). Daddy issues: paternal effects on phenotype. Cell 151, 702–708. doi: 10.1016/j.cell.2012.10.020

Rathke, C., Baarends, W. M., Awe, S., and Renkawitz-Pohl, R. (2014). Chromatin dynamics during spermiogenesis. Biochim. Biophys. Acta 1839, 155–168. doi: 10.1016/j.bbagrm.2013.08.004

Reuter, M., Berninger, P., Chuma, S., Shah, H., Hosokawa, M., Funaya, C., et al. (2011). Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. Nature 480, 264–267. doi: 10.1038/nature10672

Romero, Y., Meikar, O., Papaioannou, M. D., Conne, B., Grey, C., Weier, M., etal. (2011). Dicer1 depletion in male germ cells leads to infertility due to cumulative meiotic and spermiogenic defects. PLoS ONE 6:e25241. doi: 10.1371/journal.pone.0025241

Samans, B., Yang, Y., Krebs, S., Sarode, G. V., Blum, H., Reichenbach, M., etal. (2014). Uniformity of nucleosome preservation pattern in mammalian sperm and its connection to repetitive DNA elements. Dev. Cell 30, 23–35. doi: 10.1016/j.devcel.2014.05.023

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., et al. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol. Cell 48, 849–862. doi: 10.1016/j.molcel.2012.11.001

Sendler, E., Johnson, G. D., Mao, S., Goodrich, R. J., Diamond, M. P., Hauser, R., et al. (2013). Stability, delivery and functions of human sperm RNAs at fertilization. Nucleic Acids Res. 41, 4104–4117. doi: 10.1093/nar/gkt132

Siomi, M. C., Sato, K., Pezic, D., and Aravin, A. A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. 12, 246–258. doi: 10.1038/nrm3089

Smallwood, S. A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., etal. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat. Genet. 43, 811–814. doi: 10.1038/ng.864

Smith, Z. D., Chan, M. M., Mikkelsen, T. S., Gu, H., Gnirke, A., Regev, A., et al. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484, 339–344. doi: 10.1038/nature10960

Soboleva, T. A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G. A., and Tremethick, D. J. (2012). A unique H2A histone variant occupies the transcriptional start site of active genes. Nat. Struct. Mol. Biol. 19, 25–30. doi: 10.1038/nsmb.2161

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., etal. (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. Cell Rep. 3, 2179–2190. doi: 10.1016/j.celrep.2013. 05.031

Thomson, T., and Lin, H. (2009). The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. Annu. Rev. Cell Dev. Biol. 25, 355–376. doi: 10.1146/annurev.cellbio.24.110707.175327

Torres, A. G., Batlle, E., and Ribas De Pouplana, L. (2014). Role of tRNA modifications in human diseases. Trends Mol. Med. 20, 306–314. doi: 10.1016/j.molmed.2014.01.008

van der Heijden, G. W., Derijck, A. A., Ramos, L., Giele, M., Van Der Vlag, J., and De Boer, P. (2006). Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. Dev. Biol. 298, 458–469. doi: 10.1016/j.ydbio.2006.06.051

van der Heijden, G. W., Ramos, L., Baart, E. B., Van Den Berg, I. M., Derijck, A. A., Van Der Vlag, J., et al. (2008). Sperm-derived

histones contribute to zygotic chromatin in humans. BMC Dev. Biol. 8:34. doi: 10.1186/1471-213X-8-34

Vavouri, T., and Lehner, B. (2011). Chromatin organization in sperm may be the major functional consequence of base composition variation in the human genome. PLoS Genet. 7:e1002036. doi: 10.1371/journal.pgen.1002036

Wagner, K. D., Wagner, N., Ghanbarian, H., Grandjean, V., Gounon, P., Cuzin, F., et al. (2008). RNA induction and inheritance of epigenetic cardiac hypertrophy in the mouse. Dev. Cell 14, 962–969. doi: 10.1016/j.devcel.2008.03.009

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., et al. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev. 20, 1732–1743. doi: 10.1101/gad.14 25706

Watanabe, T., Tomizawa, S., Mitsuya, K., Totoki, Y., Yamamoto, Y., KuramochiMiyagawa, S., et al. (2011). Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. Science 332, 848–852. doi: 10.1126/science.1203919

Wu, Q., Song, R., Ortogero, N., Zheng, H., Evanoff, R., Small, C. L., et al. (2012). The RNase III enzyme DROSHA is essential for microRNA production and spermatogenesis. J. Biol. Chem. 287, 25173–25190. doi: 10.1074/jbc.M112.3 62053

Yatsenko, A. N., Roy, A., Chen, R., Ma, L., Murthy, L. J., Yan, W., et al. (2006). Noninvasive genetic diagnosis of male infertility using spermatozoal RNA: KLHL10 mutations in oligozoospermic patients impair homodimerization. Hum. Mol. Genet. 15, 3411–3419. doi: 10.1093/hmg/ddl417

Objectives

The general objective of this thesis is to elucidate the role of repetitive elements and heterochromatin in transmitting epigenetic information to the next generation. We want to disentangle how genetic elements (focusing on how transposons affect small RNAs and heterochromatin) and environmental perturbations (focusing on the effect that physiological stimuli have on small RNAs and heterochromatin) modify the germline epigenome in a way that can transmit epigenetic information between generations in animals. Specifically, we can summarize the main goals as follows:

Chapters 1 & 2

1. Study inter-individual variation in small RNA expression in mice.
2. Test the link between polymorphic repetitive elements and variation in small RNA expression.
3. Find molecular mechanisms and proteins involved in variation of small RNA expression.
4. Test whether repetitive elements have any consequence on gene expression in the next generation.

Chapter 3, 4 & 5

5. Identify genes of epigenetic pathways involved in the transmission of epigenetic information.
6. Describe genome-wide changes due to inherited epigenetic alterations.
7. Test the role of heterochromatin in the transmission and maintenance of differential expression in the next generation.
8. Test the role of repetitive elements in the transmission and maintenance of differential expression in the next generation.

Methodology

Chapter 1: Endogenous retrovirus insertions switch genes into piRNA-producing loci in mouse

Mouse tissue isolation and RNA extraction

ICR mice (ICR-CD1, Envigo) were maintained and used according to the guidelines of the Universitat de Barcelona Animal Care and Use Committee. Animals were maintained in a 12-hour light-dark cycle, under constant conditions of temperature and humidity. Adult 2 month-old mice were sacrificed under a CO2 atmosphere. Testicles were rapidly dissected, snap-frozen in liquid nitrogen and stored at -80 C. Total RNA was extracted from previously frozen testes of ICR mice using TRI Reagent® (Sigma-Aldrich, Madrid, Spain).

For genotyping, DNA extraction from mouse liver tissue was performed using the Maxwell 16 Tissue DNA Purification kit (Promega). 20 µL PCR reactions were performed with 50 ng of genomic DNA using the Phusion High Fidelity DNA Polymerase (2 U/µL) (Life Technologies) following manufacturer's indications. Specifically, 0,5 µL of 10 µM Forward primer (5' TACTAATTCCAGACCTCTCTCC 3') and Reverse primer (5' GCACTGTAGAGTCGACTGGTGC 3') were used together with 0,4 µL 10 mM dNTPs and 0,4 µL of Phusion Polymerase. PCR conditions were as follows: an activation step at 98ºC for 3'; 30 x 3-step cycles of denaturing at 98ºC for 10'', annealing at 61.2ºC for 20'' and extension at 72ºC for 4' 15''; followed by a final step at 72ºC for 5'. Amplicons were run in 0.8% agarose gels stained with SYBR safe (Life Technologies). Gel pictures were taken with Molecular Imager® Gel Doc™ XR+ imaging system (BioRad).

For the isolation of spermatogonial RNA, C57BL/6 and Cast/EiJ mice were obtained from The Jackson Laboratories and kept in the SPF animal facility of Max Planck Institute of Immunobiology and Epigenetics until sacrifice. In order to isolate spermatogonia from mice, testes were dissected and digested according to the protocol by Liao et al., Bio. Protocol (2016), with minor modifications. Briefly, we euthanized 6 weeks old mice with CO_2 and quickly dissected testes,

removed the tunica albuginea and loosened the seminiferous tubules. We then digested these tissues with 1 mg/ml collagenase IV (Worthington, LS004189) in DMEM (Gibco, 31966-024) supplemented with 10% FBS, 100 U/ml penicillin-streptomycin (Gibco, 15140-122), 250 ng/ml fungizone (Gibco, 15290-018) and 50 µg/ml gentamycin (Serva, 4799.01) in a petri dish at 37°C over a Thermoblock, shaking at 600 rpm for 30 minutes. The reaction continued for another 10 minutes after the addition of 0.25% tripsin EDTA (Sigma, T4849) at 37°C and 600 rpm. We homogenized the digested tissues by pipetting up and we washed the solution with a double amount of PBS (Gibco, 14190-094) supplemented with 10% FBS. Pieces of remaining, undigested tissues were filtered with a 40 µm strainer (BD Falcon, 352340). The filtered solution was then centrifuged at 300 g for 10 minutes at 4°C. We removed the supernatant and then resuspended the pellet in 200 µl of FACS buffer (PBS supplemented with 5% BSA and 5 mM EDTA) supplemented with 1U/µl SUPERase.in (Invitrogen, AM2696). Spermatogonia were sorted according to (Kanatsu-Shinohara et al. 2011) for the expression of CD9 (eBioscience, 17-0091-82, 1µg) and Epcam (eBioscience, 0.125 µg). Sorted cells were centrifuged at 300g for 10 minutes at 4°C and resuspended in 1 ml of TRIzol (Invitrogen, 15596018). Spermatogonial RNA was purified according to the standard TRIzol protocol and contaminant genomic DNA was digested using the DNA-free kit (Invitrogen, AM1906).

Small RNA-seq library preparation and analysis of data

Libraries were prepared with TruSeq small RNA from illumina with extended range of size selection. Pippin prep was used for automated pooled library size selection. Libraries were indexed using Illumina barcodes and sequenced using a HiSeq2500 (Illumina) as single 50nt reads. Small RNA libraries corresponding to samples from inbred strains were sequenced as a single pool on two lanes and the resulting data (all showing very high correlation between lanes) were merged for analysis.

We trimmed small RNA reads of the adaptor using cutadapt (Martin 2011) v1.9.1 and mapped them to the mouse genome (version mm10) using bowtie (Langmead et al. 2009) v1.1.2 with the options –M 1 --best --strata -v 1 to get the best alignment with up to 1 mismatch. To

identify piRNAs produced by genic transcripts, we annotated reads mapping to genes using featureCounts v1.5.1 (Liao, Smyth, and Shi 2014) with the –Q 1 option to remove multi-mapping reads from bowtie and the -s 1 option to count only reads on the same strand as the gene. For total RNA or mRNA data we mapped reads to the genome using TopHat2 v2.1.0 (Kim et al. 2013). We annotated reads using featureCounts v1.5.1 (Liao, Smyth, and Shi 2014) with the -s 1 option to count only reads on the same strand, the -t exon option to count only reads on exons and the -B and -p options to count only templates that had both pairs aligned. We mapped Chip-Seq reads using bowtie2 (Langmead and Salzberg 2012). We scaled the abundance of reads of small RNAs and Chip-seq to reads per million of uniquely mapped reads (RPM). We used the Hartigan's dip test for unimodality (Hartigan and Hartigan 1985) to find piRNA clusters with a multimodal pattern of expression across all ICR samples. For the analysis of allele-specific expression and chromatin from hybrid mouse strains we used SNPSplit (Krueger and Andrews 2016) to quantify each allele separately. Briefly, using SNPSplit we masked the mouse genome assembly changing all the SNP positions from the Sanger Institute Mouse Genomes Project to Ns. Reads overlapping such SNPs get assigned to the corresponding strain according to the genotype. We used the NCBI RefSeq gene annotation from the GRCm38 (mm10) genome assembly. For the identification of genic IAP elements we used the transposable elements annotation across 18 mouse strains from (Nellåker et al. 2012). The coordinates of 214 known mouse piRNA producing loci were retrieved from (Li et al. 2013) and correspond to inbred mouse strain C57BL/6.

Chapter 2: An IAP element drives parent-of-origin dependent gene expression

I mapped mRNA data reads to the genome using TopHat2 v2.1.0 (Kim et al. 2013). I annotated reads using featureCounts v1.5.1 (Liao, Smyth, and Shi 2014) with the -s 1 option to count only reads on the same strand, the -t exon option to count only reads on exons and the -B and -p options to count only templates that had both pairs aligned. I scaled the abundance of reads to reads per million uniquely mapped reads (RPM). For the analysis of allele-specific expression and chromatin from hybrid mouse strains we used SNPSplit (Krueger and Andrews 2016) to quantify each allele separately. Briefly, using SNPSplit we masked the mouse genome assembly changing all the SNP positions from the Sanger Institute Mouse Genomes Project to Ns. Reads overlapping such SNPs get assigned to the corresponding strain according to the genotype. We used the NCBI RefSeq gene annotation from the GRCm38 (mm10) genome assembly. For the identification of genic IAP elements we used the transposable elements annotation across 18 mouse strains from (Nellåker et al. 2012).

Chapter 3: Paternal diet defines offspring chromatin state and intergenerational obesity

Reads were mapped using TopHat v2.0.8, with -G option against the Drosophila mel- anogaster genome (assembly BDGP5, Ensembl release 69). Gene expression values and significantly differentially expressed genes were calculated using Cuffdiff v2.1.1 with upper-quartile normalization and weighting multimapping reads (-N -u options). Gene set enrichment analysis used GSEA 2.0 or GSEAPreranked with default parameters. Enrichment plots used the Cytoscape plugin Enrich- ment Map. Analysis of the five chromatin colors used BedTools (2.16.2). For microarray analyses, normalized probe values from the authors were mapped using Ensembl Biomart, and differential analysis against corresponding wild-types were performed using limma in R. Statistically significant was adjusted p value < 0.05 and fold change > 2. Enrichment of chromatin and insulator ChIP-seq data sets from modENCODE used deep- Tools 1.5.8.1 (Ramirez et al., 2014). Equivalently expressed gene sets were considered as the mean signal of the two genes ranked above and below each gene of interest. Distance to insulators was calculated using BedTools (2.16.2).

Chapter 4: Transgenerational transmission of environmental information in C. elegans

Reads were mapped using TopHat v2.1.0 (32) with the options --no-coverage-search -i 10 -I40000 -g 20 to include multi-mapping at all the possible hits against the C. elegans genome (assembly WS215) from WormBase. Read counting at different genomic features was performed using featureCounts v1.5.1 (33) with the option -s 2 -B -p and the -M --fraction option. Multi mapping reads contribute as fractional counts towards the loci they map to (a read mapped x times adds 1/x to each locus). Each copy of a repeat was annotated separately. Re-analysis of the data using uniquely mapping reads showed very similar results. Collapsing repeats by subfamily led to the same conclusions. We used the C. elegans genome annotation from Ensembl release 70 and the RepeatMakser (http://www.repeatmasker.org) annotation from the UCSC genome browser. Pseudogenes overlapping repeats and repeats overlapping exons were removed. We only considered genes and repeats with a median of one or more reads (present in at least half the samples). Data scaling, normalization and tests for differential expression were performed with DESeq2 version 1.8.1 (34). DESeq2 applies a shrinkage, or regularization method, on log2 fold changes and these are the values plotted in the figures. In addition, we processed the data (using the same filtering of genes with low read counts) using the standard limma pipeline to estimate the true biological correlation using the genas function (35). For the transformation of reads to logCPM counts we used a prior of 1 and for the calculation of differential expression we used limma-trend. The linear model was then passed to the genas function to estimate the biological correlation between the contrast of set25 mutants against the wt controls at 20°C and the 25°C wt animals against the 20°C wt animals. Results from a number of different sub-setting methods of the genas function are shown in Table S4. This analysis confirmed that DNA transposons and other repeats had a strong positive correlation between the two different contrasts.

Chapter 5: Impaired DNA replication derepresses chromatin and generates a transgenerationally inherited epigenetic memory

Sequence reads were mapped using TopHat2 version 2.1.0 (41), with default parameters against a custom genome consisting of the C. elegans genome assembly WS215 from WormBase and the sequence of the transgene vector. Reads aligning to different genomic features were counted using featureCounts version 1.5.0 (42) with the option -s 2 -M --fraction to include multimapping reads and weighting them by number of matches. We used the C. elegans genome annotation from Ensembl Release 70. Data scaling, normalization, and tests for differen- tial expression were performed using DESeq2 package version 1.8.1 (43) for R 3.2.0 (R Core Team 2015). Chromatin state segmentation and description were from modENCODE (15) using the early-stage embryo chromatin. Each gene was assigned to all overlapping states.

Results

Chapter 1: Endogenous retrovirus insertions switch genes into piRNA-producing loci in mouse

Chapter 2: An IAP element drives parent-of-origin dependent gene expression

Chapter 3: Paternal diet defines offspring chromatin state and intergenerational obesity

Chapter 4: Transgenerational transmission of environmental information in C. elegans

Chapter 5: Impaired DNA replication derepresses chromatin and generates a transgenerationally inherited epigenetic memory

Chapter 1: Endogenous retrovirus insertions switch genes into piRNA-producing loci in mouse

Authors

Eduard Casas, Cristina Moreta, Judith Cebria, Ilaria Panzeri, J. Andrew Pospisilik, Josep Jimenez-Chillaron, Sonia Forcales, Tanya Vavouri

Introduction

It is estimated that 10-12% of all spontaneous mutations in mouse are caused by endogenous retrovirus (ERV) insertions that occur in the male germline (Maksakova et al. 2006; Nellåker et al. 2012). One of the most prolific families of murine endogenous retroviruses is the intracisternal A particle (IAP). IAPs are highly expressed in pre-meiotic male germ cells (A. Dupressoir and Heidmann 1996) and are responsible for approximately a third of polymorphic transposable element insertions in laboratory mouse strains (Nellåker et al. 2012). There are multiple defence mechanisms against IAPs and other transposable elements, including the germline-specific PIWI pathway (Carmell et al. 2007; A. A. Aravin et al. 2007, 2008; De Fazio et al. 2011; Reuter et al. 2011; Di Giacomo et al. 2013). In mammals and other animals, long piRNA precursor transcripts, known as piRNA clusters, are targeted by PIWI proteins and cleaved into multiple 26-31 nucleotide small RNAs (Girard et al. 2006; A. Aravin et al. 2006; Grivna et al. 2006; Brennecke et al. 2007; Lau et al. 2006). Long non-coding RNAs, transposon transcripts and protein-coding genes act as piRNA precursors. In mouse, more than half of of all known piRNA producing loci correspond to protein-coding genes (A. A. Aravin et al. 2007; Li et al. 2013). After transcription by RNA polymerase II, and in some cases splicing, precursor transcripts are exported by unknown factors to the cytoplasm, where they are processed into mature piRNAs (reviewed in (Czech and Hannon 2016)). The dynamic landscape of transposable element insertions along the genome raises the question of how it affects the identity and diversity of piRNA producing loci in a species.

Results

We set out to analyse inter-individual variation in piRNA expression, aiming to associate piRNA expression with genetic polymorphisms and ultimately gain insight into mammalian piRNA biogenesis and evolution (Fig 5A). A similar approach previously led to the discovery of new transposon-associated piRNA clusters in different Drosophila strains (Shpiz et al. 2014). To gain insight into the role of genetic polymorphism in piRNA cluster expression, we performed this study in mouse, which is the mammalian species where the PIWI pathway is best described. As a first approach, we studied inter-individual variation in piRNA expression using genetically diverse mice of the outbred strain ICR. Using small RNA sequencing, we quantified the abundance of small RNAs in adult testis of eighteen mice and looked for known mouse piRNA producing loci (Li et al. 2013) with distinct patterns of expression across individuals. To our surprise, we found a piRNA producing locus with binary expression across the population - it appears as either expressed or silent in any one individual (Fig 5B). We refer to this cluster as pi-Noct because it maps entirely within the protein-coding gene Nocturnin (*Noct*) (Li et al. 2013).

We looked for further evidence that the small RNAs mapping to *Noct* are piRNAs. As expected of piRNAs, pi-Noct small RNAs start predominantly with uridine and are on average 27nt long (Fig 5C). The average length of these small RNAs suggests they are bound to MILI, which we confirmed by analysing available data from small RNAs immunoprecipitated with the two PIWI proteins expressed in adult testes (Ding et al. 2017) (Fig 5D). With the exception of Drosophila, polymorphic piRNA clusters have so far not been identified in any other species. Although some variation in the level of piRNA abundance has been previously observed between different zebrafish strains (Kaaij et al. 2013), to our knowledge pi-Noct is the first example of a naturally occurring locus in any vertebrate genome that produces piRNAs in some individuals in a population and not in others.

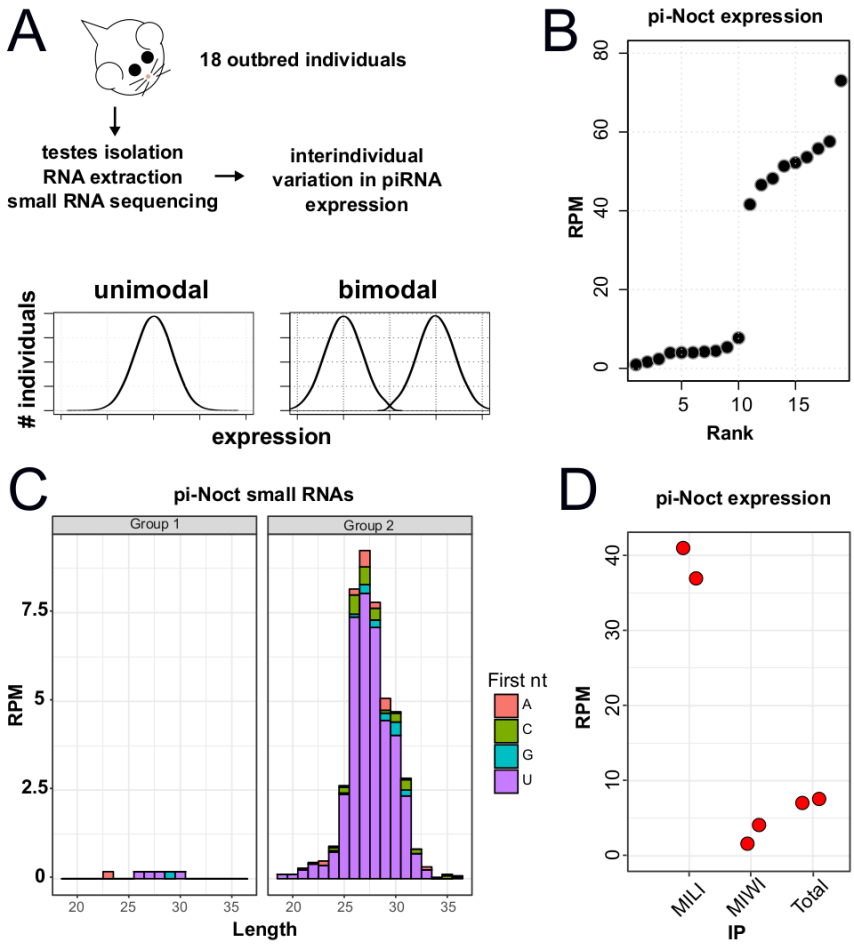


Figure 5. Inter-individual variation in piRNA expression. A) Experimental design. BpiRNA producing locus overlapping *Noct* has bimodal expression among different mice of the outbred ICR strain. B) In animals with high pi-Noct expression, small RNAs have the characteristic length and first nucleotide distribution of piRNAs. C) Single representative sample with low pi-Noct expression (left) and high pi-Noct expression (right) is shown. D) pi-Noct small RNAs are co-immunoprecipitated with MILI. MILI and MIWI co-immunoprecipitated small RNA data from testes of wild type adult C57BL/6J mice (Ding et al. 2017).

To understand the cause of binary expression of the pi-Noct piRNA cluster, we looked at its DNA sequence. In the reference mouse strain (C57BL/6), the first intron of this gene contains a 5.3kb endogenous retrovirus insertion of the IAP IΔ1 subclass (Fig 6A). The *Noct* IAP insertion is polymorphic between different laboratory inbred strains and therefore most likely occurred during the past approximately 100

years (Nellåker et al. 2012; Anne Dupressoir et al. 1999). This suggests two hypotheses: first, that the *Noct* IAP insertion is also polymorphic between different individuals of the outbred ICR strain and second, that there is an association between the IAP insertion and piRNA production from this locus. To test the first hypothesis we genotyped mice of the ICR strain (Fig 6B) and confirmed that the *Noct* IAP insertion is polymorphic within this outbred strain, with half of the animals we have interrogated being homozygous negative for the *Noct* IAP insertion (Fig 6B, blue samples, lower band). We then tested whether there is an association between the IAP insertion and pi-Noct expression in different individuals of the ICR strain. Strikingly, we found perfect agreement between genotype and pi-Noct expression - all samples producing piRNAs from *Noct* are from mice with at least one allele containing the *Noct* IAP insertion whereas all samples not producing piRNAs from *Noct* are from mice homozygous negative for the *Noct* IAP insertion (Fig 6C).

To test the link between the IAP insertion and piRNA production from *Noct* in other inbred mouse strains, we sequenced small RNAs from adult testes of two *Noct* IAP positive strains (C57BL/6 and NOD) and two *Noct* IAP negative strains (129S and C3H) (Fig 6B). In agreement with our hypothesis, piRNAs are produced from *Noct* in IAP positive strains whereas they are not produced from this locus in IAP negative strains (Fig 6C). We also analysed publicly available data from FVB/NJ (Sharma et al. 2016) and found a consistent absence of pi-Noct piRNAs in all datasets from this Noct IAP negative mouse strain (Supplementary Figure 1A, B). In summary, the IAP insertion perfectly explains piRNA production from *Noct* in the mouse germline. These results suggest that the recent insertion of an endogenous retrovirus of the IAP family in the *Noct* intron has triggered the birth of a new mouse piRNA producing locus.

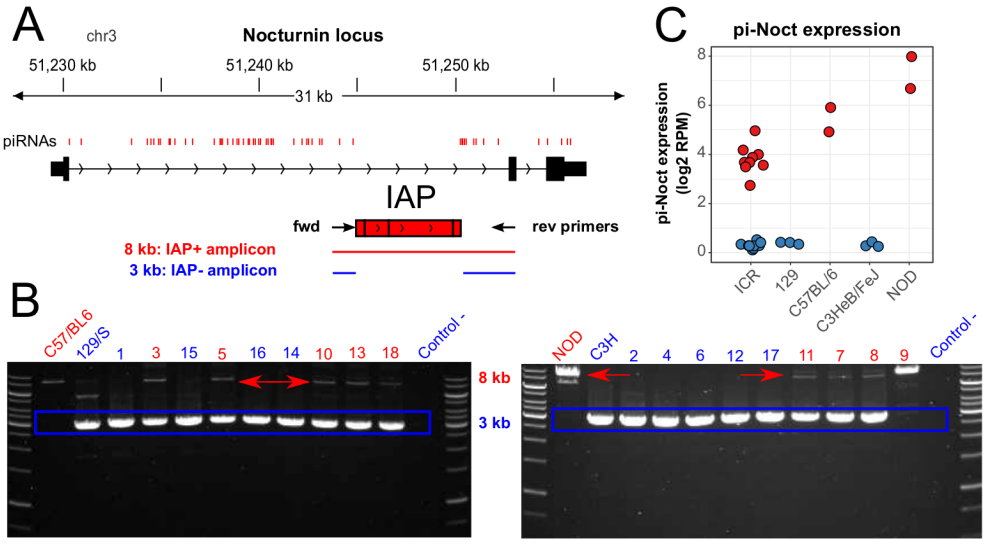


Figure 6. pi-Noct overlaps a polymorphic IAP perfectly correlated with piRNA expression. A) The pi-Noct locus. Black arrows show the location of the primers used for genotyping: the sequences amplified from the IAP positive allele (red) and the IAP negative allele (blue) are indicated. Uniquely mapping small RNAs from a representative ICR mouse sample (with high pi-Noct expression) are shown in red. All small RNAs map in the sense strand. Multimapping small RNAs on IAP are shown in Supp. Fig 1F. B) Genotyping PCR. The *Noct* IAP positive allele corresponds to an ~8kb PCR product (red rectangle) while the *Noct* IAP negative allele corresponds to an ~3kb PCR product (blue rectangle). C) Noct IAP negative individuals (shown in blue) produce no piRNAs from the pi-Noct locus, whereas *Noct* IAP positive (shown in red) individuals do.

Next, we assessed the expression of pi-Noct during spermatogenesis. We analysed available oxidised small RNA data from early postnatal development (Li et al. 2013) and found that pi-Noct is most highly expressed at 10.5 days postpartum when the synchronously developing germ cells reach the primary spermatocyte stage (Fig 7A, Supp. Fig 1E). pi-Noct expression drops in subsequent developmental stages, following the trend of the so-called pre-pachytene piRNAs (A. A. Aravin et al. 2007; Li et al. 2013) (Fig 7A, Supp. Fig 1E). Similar to other pre-pachytene piRNAs (Li et al. 2013), pi-Noct small RNAs are processed entirely from the sense strand of the gene (Fig 6A). This data suggests that the *Noct* IAP provides signals that turn a protein-coding gene into a pre-pachytene piRNA-producing locus.

IAPs can affect gene expression in multiple ways, one of which is by acting as promoters or enhancers. Thus, we asked whether piRNA production is explained by IAP-induced ectopic transcriptional activation of the gene during spermatogenesis. To address this, we analysed available steady state gene expression data from different stages of spermatogenesis from mouse hybrids carrying one *Noct* IAP positive and one *Noct* IAP negative allele (Gan et al. 2013) and from this data we quantified the expression of each allele using exonic single nucleotide variants linked to the IAP. Throughout spermatogenesis, *Noct* is among the most highly expressed genes (Fig 3B) with no evidence of the Noct IAP positive allele being more highly expressed than the ancestral, Noct IAP negative allele (Fig 7C). Similarly, we analysed the chromatin state of the *Noct* promoter using available H3K4me3 ChIP-seq data from spermatocytes of mouse hybrids (Baker et al. 2015) and found, again, that the promoter is in active state independently of the presence of the IAP (Fig 7D). These results argue that the *Noct* IAP inserted into a pre-existing germline-expressed gene and that it did not change the expression of the gene at the transcriptional level. The simplest explanation of the observed data is that the Noct IAP carries post-transcriptional regulatory signals that funnel genic transcripts to the piRNA biogenesis pathway.

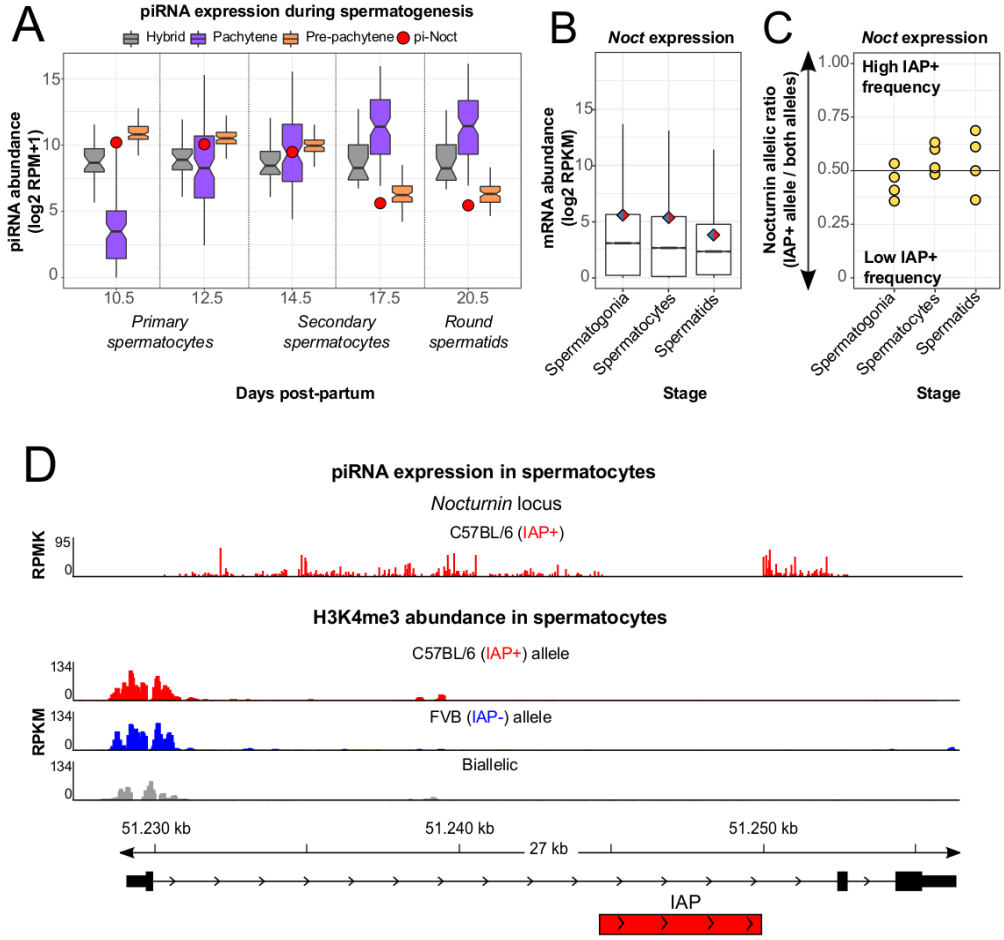


Figure 7. The IAP insertion is not required for Noct transcription during spermatogenesis. A) pi-Noct piRNA expression during spermatogenesis. Oxidised small RNA data from testes of C57BL/6 mice from (Li et al. 2013). B) N*oct* is highly expressed in spermatogonia, spermatocytes and spermatids from both the Noct IAP+ and the Noct IAP- allele (129/DBA F1 hybrid strain; data from (Gan et al. 2013)) showing that the IAP insertion is not required for *Noct* expression during spermatogenesis. C) Both alleles of *Noct* are equally expressed in a hybrid strain that contains a Noct IAP positive and a Noct IAP negative allele. D) pi-Noct expression in oxidised small RNA data from early spermatocytes of C57BL/6 mice from (Li et al. 2013) on the IAP-containing intron (top) and H3K4me3 in the promoter of *Noct* in IAP positive and IAP negative strains (Baker et al. 2015) (bottom).

To substantiate the association between piRNA production from genes and IAP insertions, we searched for other examples of genic piRNA producing loci overlapping polymorphic IAPs. We compared available spermatid small RNA data from mouse strains C57BL/6 (Gainetdinov et al. 2018) and FVB/NJ (Sharma et al. 2016) and looked for IAPs that are present in the former strain but are absent from the latter strain. We identified two genic piRNA producing loci with such polymorphic IAPs, with the IAP missing from FVB/NJ in both cases. The two genes with polymorphic IAPs are *Noct* and *Phf20*. In both cases, piRNA production is absent in FVB/NJ which is the strain missing the IAP (Supp. Fig 2A). Similarly to pi-Noct, pi-Phf20 small RNAs follow the expression pattern of pre-pachytene piRNAs and are predominantly bound to MILI (Supp. Fig 2B,C). *Phf20* is also highly expressed during spermatogenesis with no evidence of IAP-dependent expression (Supp. Fig 2D-F). In summary, we have identified two protein-coding genes that have independently switched to producing piRNAs at the pre-pachytene stage following two independent IAP insertions.

Next, we elucidated how IAP insertions cause piRNA production from genic transcripts. Both at *Noct* and *Phf20*, piRNA production is linked to an IAP insertion in an intron with the vast majority of piRNAs being processed from the IAP-containing intron (Fig 6A, 3D, Supp. Fig 2F). IAPs are extremely rare in exons presumably due to strong negative selection (Nellåker et al. 2012). We therefore focused on understanding how piRNAs are produced from genes with IAPs inserted in their introns. piRNA production from introns is unexpected because precursor transcripts are processed into piRNAs in the cytoplasm (A. A. Aravin et al. 2008) while introns are typically retained in the nucleus. Consequently, the first requirement for piRNA production from introns is export of introns or intron-retaining transcripts to the cytoplasm. As part of their life cycle, retroviruses have evolved mechanisms to transport their unspliced primary transcripts to the cytoplasm. One way that IAP retroviruses are thought to achieve this is by tethering their transcripts to the RNA export factor NXF1 (Floyd et al. 2003; Concepcion et al. 2015; Lindtner et al. 2006). We therefore reasoned that piRNA production from IAP-containing introns likely depends on NXF1. NXF1 has been implicated in euchromatic piRNA precursor transport in Drosophila (Handler et al. 2013; Dennis et al. 2016)) making it a potential evolutionarily

conserved component of the piRNA biogenesis pathway. To test this hypothesis, we turned to the inbred mouse strain Cast/EiJ which carries a natural allele of *Nxf1* previously shown to suppress phenotypes linked to IAP insertional mutations (Floyd et al. 2003; Concepcion et al. 2015).

Hypothesizing that the Cast/EiJ *Nxf1* allele blocks piRNA production from IAP-containing piRNA clusters, we sequenced and analysed spermatogonial small RNAs from Cast/EiJ mice and from C57BL/6 for comparison. First, we compared the proportion of reads mapping to piRNA clusters in the two mouse strains and found them to be very similar (Fig 8A,B). Thus the Cast *Nxf1* allele does not overtly affect piRNA production. We then looked at the two IAP-associated polymorphic piRNA clusters that we previously identified. In Cast/EiJ, there is complete absence of piRNA from the first *Noct* intron (Fig 8A, 2nd panel and Supp. Fig 1C-D). As Cast/EiJ is missing the *Noct* IAP insertion, this result can be explained by either the absence of the IAP or by the presence of the Cast/EiJ *Nxf1* allele. On the other hand, the Phf20 IAP insertion is present in Cast/EiJ, yet piRNA production from this gene is reduced in the strain carrying the mutant *Nxf1* allele (Fig 8A, 3rd panel). Loss of pi-Phf20 small RNAs in the Cast/EiJ strain is specific to the IAP-containing intron (Fig 8C) supporting the hypothesis that IAPs flag intron transcripts for nuclear export and piRNA processing in an NXF1-dependent manner.

Collectively, these data suggest that other IAP-containing introns of expressed genes (including introns not previously annotated as piRNA producing loci) produce piRNAs and that piRNA production from these loci requires the activity of a wild type *Nxf1*. We therefore compared piRNA production in IAP-containing introns against all other introns of spermatogonia expressed genes in the two mouse strains. We found that, in strain C57BL/6, IAP-containing introns produce significantly more piRNAs than other introns (Fig 8D). And that in agreement with our hypothesis, in the Cast/EiJ strain, there is suppression of piRNA production from IAP containing introns (Fig 8D). For example, gene Zfp69 contains an IAP insertion in its intron in C57BL/6 and produces piRNAs from the same intron (even though it was not included in the "gold standard" set of piRNA producing loci) (Fig 8A, 4th panel, 8E). In Cast/EiJ strain, which also carries the IAP insertion piRNA production

from the intron is lost (Fig 8B,E). We conclude that IAPs are in general associated with piRNA production from introns of protein-coding genes and that the Cast/EiJ mouse strain, which carries a mutant NXF1, has suppressed production of piRNAs from intronic IAPs.

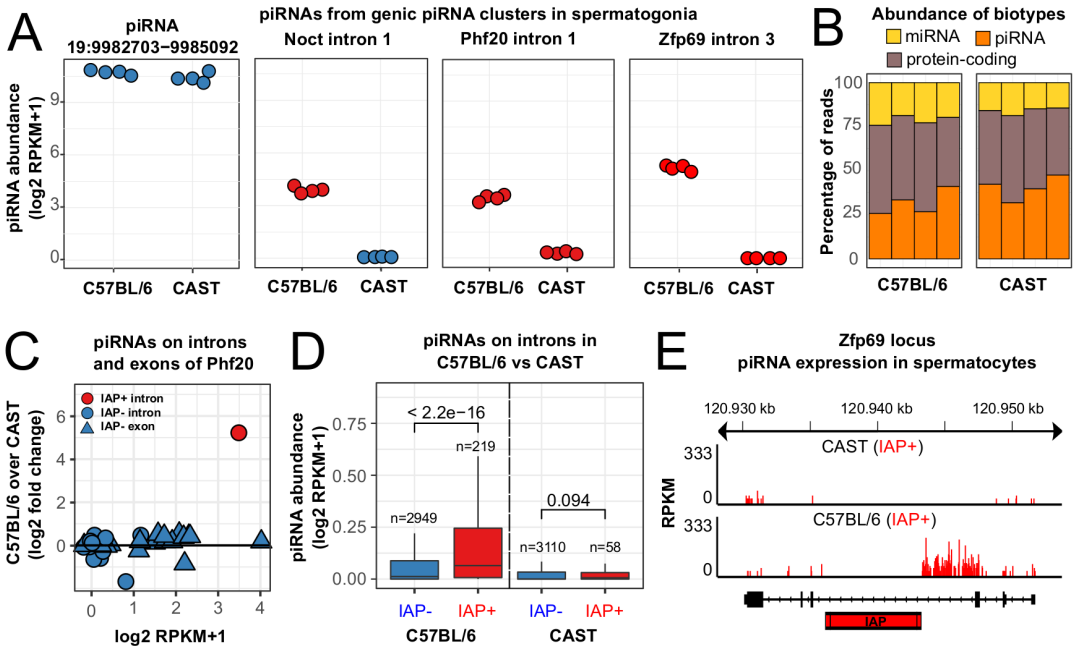


Figure 8. piRNAs from IAP-containing introns are dependent on NXF1. A) piRNAs from different clusters. The first panel shows piRNA 19:9982703−9985092, a cluster that shows no difference in expression in C57BL/6 compared to CAST. In contrast, the second panel shows Nocturnin, that is IAP- in CAST and not expressed. The third and fourth panels show genic piRNA clusters that are IAP positive in both C57BL/6 and CAST but are not expressed in CAST. B) Small RNA expression based on the biotype of the genes they come from. piRNAs are equally abundant in CAST and in C57BL/6 strains. C) piRNAs mapping to the gene body of *Phf20* separated into introns (circles) and exons (triangles). There is a big difference in piRNAs generated from the IAP-containing intron (in red) between C57BL/6 and CAST/EiJ (both Noct-IAP positive) mice while small RNAs from the rest of introns and exons remain equally abundant. D) piRNAs mapping to IAP-containing introns from expressed genes (RPKM >1). There is a significant association between piRNA expression and the presence IAPs in introns genome-wide. Cast/Eij mice do not produce piRNAs from IAP-containing introns. E) Snapshot of piRNA abundance from pi-Zfp69, showing only expression in C57BL/6 in comparison to CAST in spite of both being Zfp69 IAP positive strains.

Conclusions

To sum up, through the analysis of piRNA expression in different mouse strains we identified piRNA producing loci that are private to different individuals and strains. The presented evidence supports the idea that new piRNA producing loci emerge from germline-expressed protein-coding genes. These are triggered by the insertion of endogenous retroviruses in introns, in particular IAP transposons that have greatly expanded in laboratory mouse strains during the past 100 years. The findings presented here substantiate previous observations that protein-coding transcripts producing piRNAs from their 3' UTRs frequently contain repetitive sequences (A. A. Aravin et al. 2007). Moreover, we present how a mouse strain carrying a natural mutant allele of the RNA nuclear export factor NXF1 suppresses production of piRNAs associated with intronic IAP insertions. This suggests that a fully functional NXF1 protein is required for piRNA production from IAP-associated piRNA precursors, implicating it in the piRNA pathway in mouse for the first time (Fig 9).

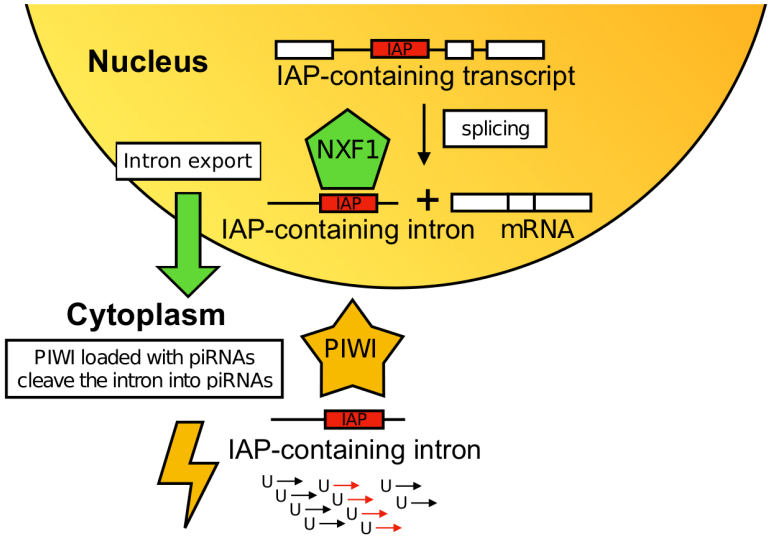
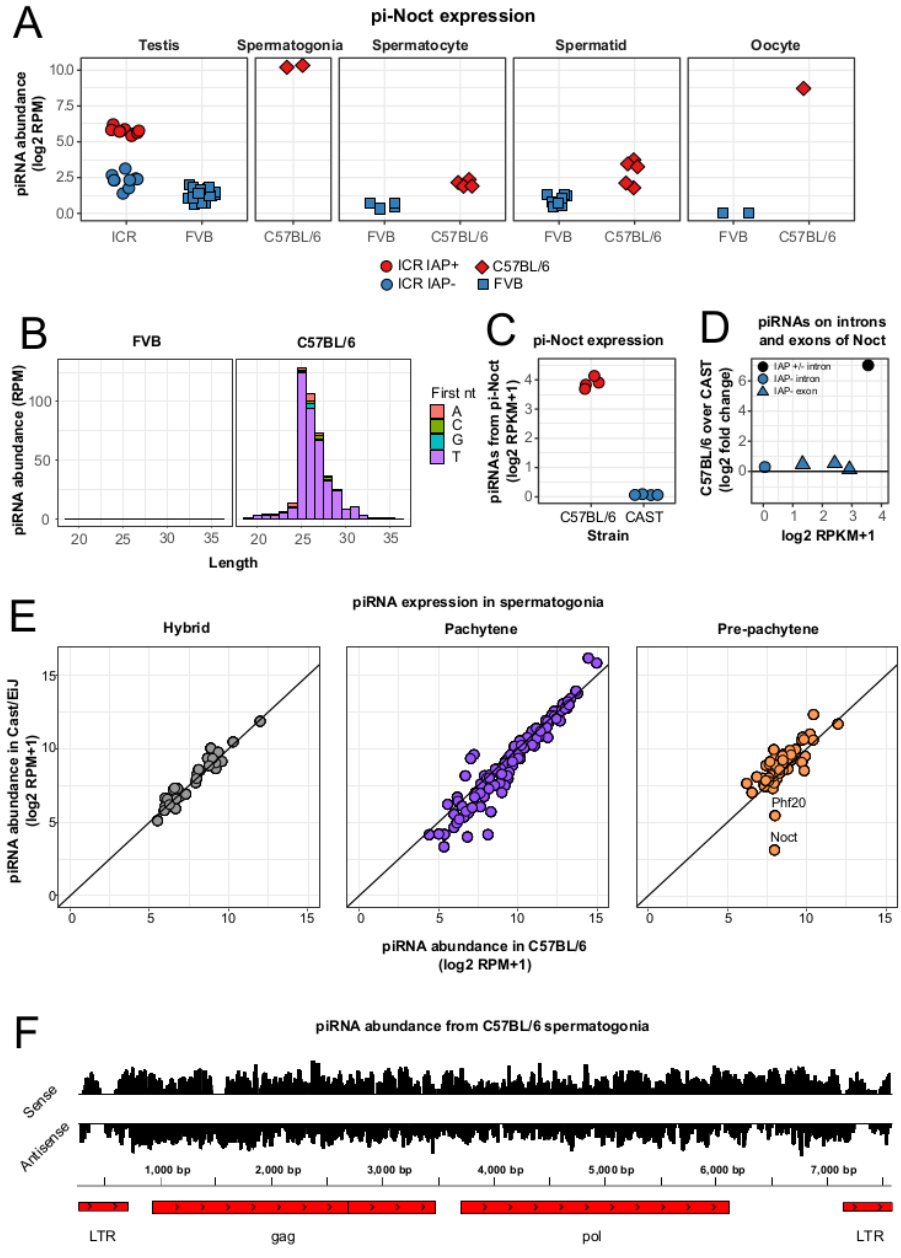
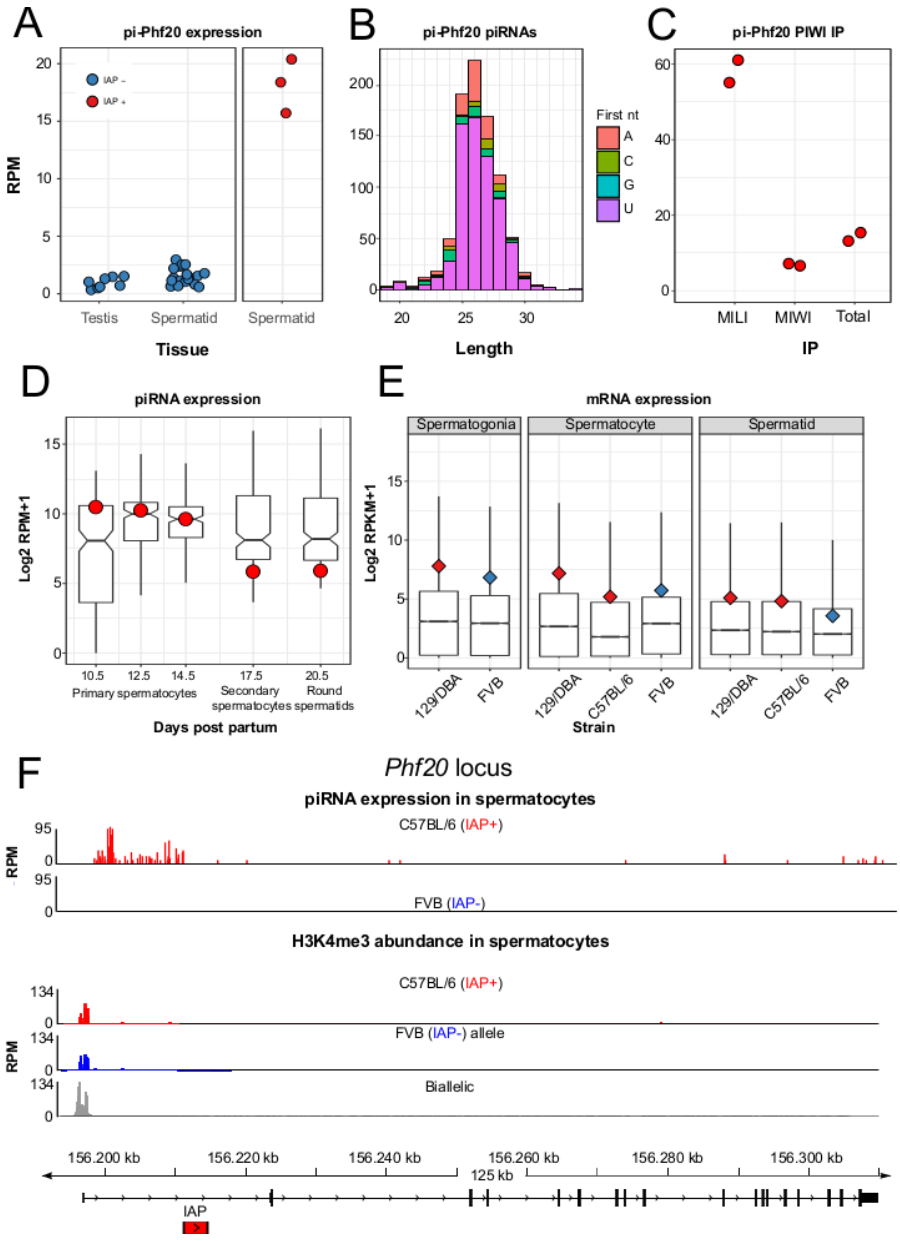


Figure 9. Mechanistic model of piRNA biogenesis. A spliced IAP-containing intron is exported to the cytoplasm by NXF1 and is recognized by PIWI as a piRNA precursor transcript, selecting it for piRNA biogenesis.

Supplementary Figures



Supplementary figure 1. A) pi-Noct is expressed in germline tissue of C57BL/6 (Noct IAP+) but not in germline tissue of FVB (Noct IAP-). B) The small RNAs spanning pi-Noct in oocytes are also piRNAs. C) piRNAs from *Noct* first intron in C57BL/6 (Noct-IAP positive) and CAST/EiJ (Noct-IAP negative). In Cast/EiJ, the Noct intron does not contain the IAP insertion and does not generate any piRNAs. D) piRNAs mapping to the gene body of *Noct* separated by introns (circles) and exons (triangles). There is a big difference in piRNAs generated from the intron (in black) that contains the IAP in C57BL/6 between such strain (Noct-IAP positive) and CAST/EiJ (Noct-IAP negative) mice, while small RNAs from the rest of introns and exons remain equally abundant. E) piRNA expression in CAST vs C57BL/6 from hybrid loci, pre-pachytene loci and pachytene loci. F) IAP elements are targeted by thousands of piRNAs both sense and antisense.



Supplementary figure 2. pi-Phf20 is another polymorphic piRNA cluster with piRNA expression linked to the IAP presence. A) The pi-Phf20 cluster is not expressed in the male germline of inbred mouse strain FVB/NJ which lacks the IAP insertion (Sharma et al. 2016) while it is expressed in C57BL/6 (Yuan et al. 2016). B) pi-Phf20 small RNAs start with a U and range between 25 and 28 nucleotides long suggesting they are likely piRNAs. C) pi-Phf20 small RNAs are also preferentially bound by MILI (data from (Ding et al. 2017)). D) pi-Phf20 piRNA expression during

spermatogenesis. Oxidised small RNA data from testes of C57BL/6 mice from (Li et al. 2013). E) Phf20 is highly expressed from both the Phf20 IAP+ and the Phf20 IAP- allele in spermatogonia, spermatocytes and spermatids IAP+/- hybrids (data from (Gan et al. 2013)) showing that the IAP insertion is not required for Phf20 expression during spermatogenesis. F) pi-Phf20 expression in small RNA data from early spermatocytes of C57BL/6 mice (Li et al. 2013) and FVB mice (Sharma et al. 2016) on the IAP-containing intron (top) and H3K4me3 in the promoter of Phf20 in IAP+ and IAP- strains (Baker et al. 2015) (bottom).

Acknowledgements

We are grateful to M Vives Pi, PJ Cardona and J Llorens for tissues from C3HeB/FeJ, 129S and NOD mice. We thank the IGTP Genomics Core Facility for small RNA library preparation, the IGTP High Performance Computing Core Facility for systems administration and the CRG Genomics Unit for small RNA sequencing. We thank Ben Lehner, Adelheid Lempdradl and Anita Öst for comments on the manuscript. EC was funded by an AGAUR FI PhD fellowship. This work was funded by Plan Estatal Grant BFU2015-70581 from the Spanish Ministry of Economy and Competitiveness and the CERCA Programme/Generalitat de Catalunya.

Author contributions

Conceived and designed the experiments: EC SF TV. Performed the experiments: EC CM JC IP SF. Analyzed the data: EC CM SF TV. Contributed reagents/materials/analysis tools: EC CM IP JC. Wrote the paper with input from all authors: EC TV. Supervised: TV AP SF JC.

Data availability

All sequencing data produced for this publication have been deposited in the NCBI Gene Expression Omnibus under accession number GSE121138.

Bibliography

Aravin, Alexei A., Ravi Sachidanandam, Deborah Bourc'his, Christopher Schaefer, Dubravka Pezic, Katalin Fejes Toth, Timothy Bestor, and Gregory J. Hannon. 2008. "A piRNA Pathway Primed by Individual Transposons Is Linked to de Novo DNA Methylation in Mice."

Molecular Cell 31 (6): 785–99.

Aravin, Alexei A., Ravi Sachidanandam, Angelique Girard, Katalin Fejes-Toth, and Gregory J. Hannon. 2007. "Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control." *Science* 316 (5825): 744–47.

Aravin, Alexei, Dimos Gaidatzis, Sébastien Pfeffer, Mariana Lagos-Quintana, Pablo Landgraf, Nicola Iovino, Patricia Morris, et al. 2006. "A Novel Class of Small RNAs Bind to MILI Protein in Mouse Testes." *Nature* 442 (7099): 203–7.

Baker, Christopher L., Shimpei Kajita, Michael Walker, Ruth L. Saxl, Narayanan Raghupathy, Kwangbom Choi, Petko M. Petkov, and Kenneth Paigen. 2015. "PRDM9 Drives Evolutionary Erosion of Hotspots in Mus Musculus through Haplotype-Specific Initiation of Meiotic Recombination." *PLoS Genetics* 11 (1): e1004916.

Brennecke, Julius, Alexei A. Aravin, Alexander Stark, Monica Dus, Manolis Kellis, Ravi Sachidanandam, and Gregory J. Hannon. 2007. "Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila." *Cell* 128 (6): 1089–1103.

Carmell, Michelle A., Angélique Girard, Henk J. G. van de Kant, Deborah Bourc'his, Timothy H. Bestor, Dirk G. de Rooij, and Gregory J. Hannon. 2007. "MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline." *Developmental Cell* 12 (4): 503–14.

Concepcion, Dorothy, Kevin D. Ross, Kasey R. Hutt, Gene W. Yeo, and Bruce A. Hamilton. 2015. "Nxf1 Natural Variant E610G Is a Semi-Dominant Suppressor of IAP-Induced RNA Processing Defects." *PLoS Genetics* 11 (4): e1005123.

Czech, Benjamin, and Gregory J. Hannon. 2016. "One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing." *Trends in Biochemical Sciences* 41 (4): 324–37.

De Fazio, Serena, Nenad Bartonicek, Monica Di Giacomo, Cei Abreu-Goodger, Aditya Sankar, Charlotta Funaya, Claude Antony, Pedro N. Moreira, Anton J. Enright, and Dónal O'Carroll. 2011. "The Endonuclease Activity of Mili Fuels piRNA Amplification That Silences LINE1 Elements." *Nature* 480 (7376): 259–63.

Dennis, Cynthia, Emilie Brasset, Arpita Sarkar, and Chantal Vaury. 2016. "Export of piRNA Precursors by EJC Triggers Assembly of Cytoplasmic Yb-Body in Drosophila." *Nature Communications* 7 (December): 13739.

Di Giacomo, Monica, Stefano Comazzetto, Harpreet Saini, Serena De Fazio, Claudia Carrieri, Marcos Morgan, Lina Vasiliauskaite, Vladimir Benes, Anton J. Enright, and Dónal O'Carroll. 2013. "Multiple Epigenetic Mechanisms and the piRNA Pathway Enforce LINE1 Silencing during Adult Spermatogenesis." *Molecular Cell* 50 (4): 601–8.

Ding, Deqiang, Jiali Liu, Kunzhe Dong, Uros Midic, Rex A. Hess, Huirong Xie, Elena Y. Demireva, and Chen Chen. 2017. "PNLDC1 Is Essential for piRNA 3' End Trimming and Transposon Silencing during Spermatogenesis in Mice." *Nature Communications* 8 (1). https://doi.org/10.1038/s41467-017-00854-4.

Dupressoir, A., and T. Heidmann. 1996. "Germ Line-Specific Expression of Intracisternal A-Particle Retrotransposons in Transgenic Mice." *Molecular and Cellular Biology* 16 (8): 4495–4503.

Dupressoir, Anne, Willy Barbot, Marie-Paule Loireau, and Thierry Heidmann. 1999. "Characterization of a Mammalian Gene Related to the Yeast CCR4 General Transcription Factor and Revealed by Transposon Insertion." *The Journal of Biological Chemistry* 274 (43): 31068–75.

Ernst, Christina, Duncan T. Odom, and Claudia Kutter. 2017. "The Emergence of piRNAs against Transposon Invasion to Preserve Mammalian Genome Integrity." *Nature Communications* 8 (1): 1411.

Floyd, Jennifer A., David A. Gold, Dorothy Concepcion, Tiffany H. Poon, Xiaobo Wang, Elizabeth Keithley, Dan Chen, et al. 2003. "A Natural Allele of Nxf1 Suppresses Retrovirus Insertional Mutations." *Nature Genetics* 35 (3): 221–28.

Gainetdinov, Ildar, Cansu Colpan, Amena Arif, Katharine Cecchini, and Phillip D. Zamore. 2018. "A Single Mechanism of Biogenesis, Initiated and Directed by PIWI Proteins, Explains piRNA Production in Most Animals." *Molecular Cell* 71 (5): 775–90.e5.

Gan, Haiyun, Lu Wen, Shangying Liao, Xiwen Lin, Tingting Ma, Jun Liu, Chun-Xiao Song, et al. 2013. "Dynamics of 5-Hydroxymethylcytosine during Mouse Spermatogenesis." *Nature Communications* 4: 1995.

Girard, Angélique, Ravi Sachidanandam, Gregory J. Hannon, and Michelle A. Carmell. 2006. “A Germline-Specific Class of Small RNAs Binds Mammalian Piwi Proteins.” *Nature* 442 (7099): 199–202.

Grivna, Shane T., Ergin Beyret, Zhong Wang, and Haifan Lin. 2006. “A Novel Class of Small RNAs in Mouse Spermatogenic Cells.” *Genes & Development* 20 (13): 1709–14.

Handler, Dominik, Katharina Meixner, Manfred Pizka, Kathrin Lauss, Christopher Schmied, Franz Sebastian Gruber, and Julius Brennecke. 2013. “The Genetic Makeup of the Drosophila piRNA Pathway.” *Molecular Cell* 50 (5): 762–77.

Hartigan, J. A., and P. M. Hartigan. 1985. “The Dip Test of Unimodality.” *Annals of Statistics* 13 (1): 70–84.

Kaaij, Lucas J. T., Suzanne W. Hoogstrate, Eugene Berezikov, and René F. Ketting. 2013. “piRNA Dynamics in Divergent Zebrafish Strains Reveal Long-Lasting Maternal Influence on Zygotic piRNA Profiles.” *RNA* 19 (3): 345–56.

Kanatsu-Shinohara, Mito, Seiji Takashima, Kei Ishii, and Takashi Shinohara. 2011. “Dynamic Changes in EPCAM Expression during Spermatogonial Stem Cell Differentiation in the Mouse Testis.” *PloS One* 6 (8): e23663.

Kim, Daehwan, Geo Pertea, Cole Trapnell, Harold Pimentel, Ryan Kelley, and Steven L. Salzberg. 2013. “TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions.” *Genome Biology* 14 (4): R36.

Krueger, Felix, and Simon R. Andrews. 2016. “SNPsplit: Allele-Specific Splitting of Alignments between Genomes with Known SNP Genotypes.” *F1000Research* 5 (June): 1479.

Langmead, Ben, and Steven L. Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): 357–59.

Langmead, Ben, Cole Trapnell, Mihai Pop, and Steven L. Salzberg. 2009. “Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome.” *Genome Biology* 10 (3): R25.

Lau, N. C., A. G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakano, D. P. Bartel, and R. E. Kingston. 2006. “Characterization of the piRNA Complex from Rat Testes.” *Science* 313 (5785): 363–67.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. “featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features.” *Bioinformatics* 30 (7): 923–30.

Lindtner, Susan, Andrei S. Zolotukhin, Hiroaki Uranishi, Jenifer Bear, Viraj Kulkarni, Sergey Smulevitch, Martina Samiotaki, George Panayotou, Barbara K. Felber, and George N. Pavlakis. 2006. “RNA-Binding Motif Protein 15 Binds to the RNA Transport Element RTE and Provides a Direct Link to the NXF1 Export Pathway.” *The Journal of Biological Chemistry* 281 (48): 36915–28.

Li, Xin Zhiguo, Christian K. Roy, Xianjun Dong, Ewelina Bolcun-Filas, Jie Wang, Bo W. Han, Jia Xu, et al. 2013. “An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes.” *Molecular Cell* 50 (1): 67–81.

Maksakova, Irina A., Mark T. Romanish, Liane Gagnier, Catherine A. Dunn, Louie N. van de Lagemaat, and Dixie L. Mager. 2006. “Retroviral Elements and Their Hosts: Insertional Mutagenesis in the Mouse Germ Line.” *PLoS Genetics* 2 (1): e2.

Martin, Marcel. 2011. “Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads.” *EMBnet.journal* 17 (1): 10.

Nellåker, Christoffer, Thomas M. Keane, Binnaz Yalcin, Kim Wong, Avigail Agam, T. Grant Belgard, Jonathan Flint, David J. Adams, Wayne N. Frankel, and Chris P. Ponting. 2012. “The Genomic Landscape Shaped by Selection on Transposable Elements across 18 Mouse Strains.” *Genome Biology* 13 (6): R45.

Reuter, Michael, Philipp Berninger, Shinichiro Chuma, Hardik Shah, Mihoko Hosokawa, Charlotta Funaya, Claude Antony, Ravi Sachidanandam, and Ramesh S. Pillai. 2011. “Miwi Catalysis Is Required for piRNA Amplification-Independent LINE1 Transposon Silencing.” *Nature* 480 (7376): 264–67.

Sharma, Upasna, Colin C. Conine, Jeremy M. Shea, Ana Boskovic, Alan G. Derr, Xin Y. Bing, Clemence Belleannee, et al. 2016. “Biogenesis and Function of tRNA Fragments during Sperm Maturation and Fertilization in Mammals.” *Science* 351 (6271): 391–96.

Shpiz, Sergey, Sergei Ryazansky, Ivan Olovnikov, Yuri Abramov, and Alla Kalmykova. 2014. “Euchromatic Transposon Insertions Trigger Production of Novel Pi- and Endo-siRNAs at the Target Sites in the

Drosophila Germline." *PLoS Genetics* 10 (2): e1004138.

Yuan, Shuiqiao, Chong Tang, Andrew Schuster, Ying Zhang, Huili Zheng, and Wei Yan. 2016. "Paternal Pachytene piRNAs Are Not Required for Fertilization, Embryonic Development and Sperm-Mediated Epigenetic Inheritance in Mice." *Environmental Epigenetics* 2 (4): dvw021.

Chapter 2: An IAP element drives parent-of-origin dependent gene expression

Authors

Eduard Casas and Tanya Vavouri

Introduction

Imprinting is an epigenetic mechanism that ensures that some genes are only expressed from the maternal or the paternal allele. There are notable imprinting differences in mammals yet it is unknown how imprinted genes evolve.

ERVs are enriched around imprinted genes (Luedi, Hartemink, and Jirtle 2005; Ke et al. 2002). IAPs are the youngest type of ERVs and they are still active in the mouse genome. ERVs regulate gene expression in embryos, for instance by driving H3K4me1- and H3K27ac-defined enhancers, (Chuong et al. 2013) and by leading to a reduction of non-terminated transcripts when inherited from the father but not from the mother at the *Slc15a2* locus (Li et al. 2012).

Interestingly, the PIWI pathway is a relevant player in imprinting in mice. For instance, its components are required for de novo methylation of the ICR of Rasgrf1 (Watanabe et al. 2011). In this case, piRNAs contribute to the methylation of the ICR leading to Rasgrf1 expression. My previous observation (reported on Chapter 1) links a polymorphic IAP to piRNA expression of a protein-coding gene. Based on the role of piRNAs in methylating ICRs and the described role of ERVs in regulating gene expression in embryos, I wondered whether *Noct* could be an example of gene regulated by ERV during embryogenesis. The hypothesis is that a recent IAP insertion leads to the emergence of imprinted expression from Nocturnin.

Results

Imprinted genes are usually discovered using F1 hybrids of reciprocal crosses of two different strains (Xie et al. 2012). In contrast to allele-specific expression, where two different alleles of the same gene are expressed at different levels, imprinted genes have the same allele differently expressed dependent on the parent of origin.

Hence, to identify potential regulatory effects of the IAP element during embryogenesis I seek to compare the expression dynamics of Noct IAP+ and IAP- alleles in two different contexts: when inherited from the father and when inherited from the mother. To do so, I analyze single-embryo mRNA sequencing datasets from F1 hybrids of IAP+ and IAP- inbred mouse strains. Through a catalog of SNPs from the Mouse Genomes Project (Keane et al. 2011) I am able to separate reads coming from the maternal or paternal strain, hence to accurately profile the expression of each allele.

Using data from single-cell embryo sequencing of CAST (IAP- strain) mothers and C57BL/6 (IAP+ strain) fathers (Deng et al. 2014) we show that the maternal allele is inherited from the oocytes but the RNA abundance is diluted after each cell division. However, during development the paternal allele increases in abundance, suggesting it is being specifically transcribed (Fig 10A). I confirmed that there is transcription of the Nocturnin locus at 2-cell stage is using androgenetic (AG) and gynogenetic (GG) embryos treated with α-amanitin-treated. A-amanitin is an inhibitor of RNA polymerase II and III, hence treated embryos do not have de novo transcription. AG and GG embryos are developed from a starting zygote with 2 nuclei from either the father of the mother respectively. Data from this experiment (Inoue et al. 2017) allows us to identify parental-specific gene expression by having expression of two paternal alleles in androgenetic embryos and of two maternal alleles in gynogenetic embryos. Consistent with the onset of transcription at 2-cell stage, α-amanitin-treated embryos show a depletion of Noct at 2-cell stage (Fig 10B), but we observe no differences between AG and GG embryos.

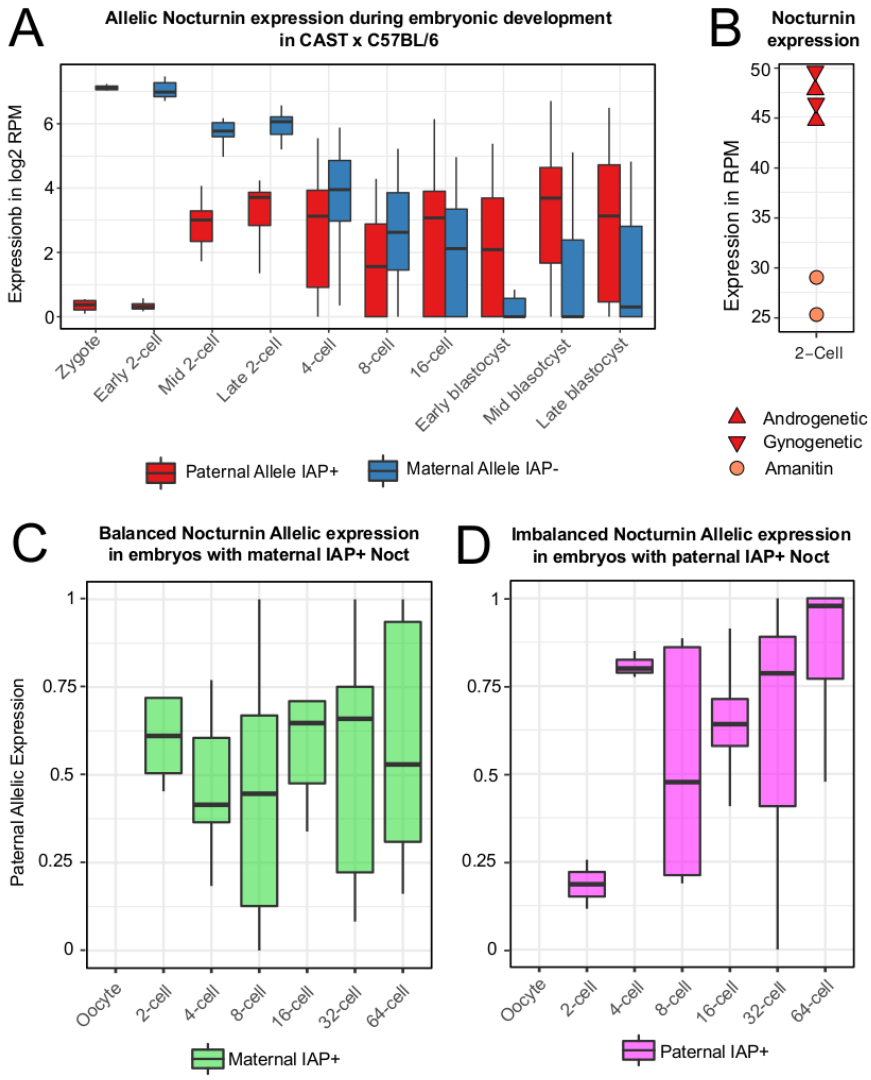


Figure 10. Allelic expression of Nocturnin gene depending on the IAP presence and the parent of origin. A) In the CAST (mother, IAP-) x C57BL/6 (father, IAP+) cross, the maternal allele (IAP-, blue boxplots) is highly abundant at the zygote stage due to the maternal load in oocytes and decreases during development. On the contrary, on the zygote the paternal allele (IAP+, red boxplots) is negligible due to the lesser sperm but is actively transcribed during development. B) α-amanitin treated embryos show a lower amount of *Noct* transcript due to hampered transcription. C) and D). Y-axis shows ratio. When the IAP+ allele is inherited from the mother both alleles have the same level of expression. When the IAP+ allele is inherited from the father it dominates the expression level of the *Noct* gene.

At the blastocyst stage the paternal allele dominates the expression of the gene. This suggests that either the increased expression is due to the paternal allele containing the IAP or that *Noct* is a novel paternally expressed gene. I test so using a dataset (Borensztein et al. 2017) that provides an excellent resource of mRNA expression of single embryos coming from reciprocal crosses of C57BL/6 x CAST. Allele-specific expression profiling of these hybrids shows that in blastocyst stage the IAP+ allele is expressed at the same level as the IAP- when inherited from the mother (Fig 10C) but it is much more expressed when it is inherited from the father (Fig 10D). In summary, the paternal allele starts being transcribed at late 2-cell stage independently of the IAP variation but when it is inherited from the father becomes the only active allele. Thus, it behaves in a parent-of-origin specific manner consistent with imprinted loci.

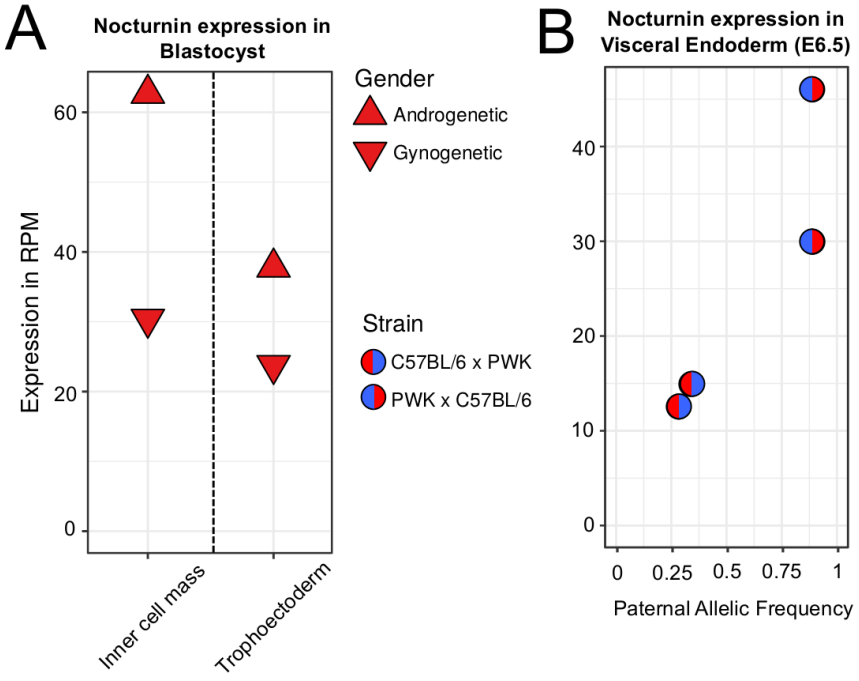


Figure 11. *Nocturnin* is imprinted when the IAP+ allele is inherited from the mother. A) In an IAP+ strain, *Nocturnin* is more expressed in two different tissues at the blastocyst stage in androgenetic embryos (those with 2 paternal nuclei) compared to gynogenetic embryos (those with 2 maternal nuclei). B) In visceral endoderm at embryonic day 6.5, the IAP+ allele is always the most transcribed. Also, when the IAP+ allele is inherited from the father it is much more expressed.

To further show this phenomenon I use AG versus GG blastocysts (Fig 11A) and visceral endoderm from 6.5 day-old embryos (Fig 11B) (Inoue et al. 2017). Here, we have embryos further developed from two paternal (AG) or maternal (GG) pronuclei. At this stage, imprinting expression should be clearly biased. Interestingly, we find that the IAP+ allele is more highly expressed when inherited from the father consistent with previous results and proving that the gene is imprinted with epigenetic modifications associated with increased expression.

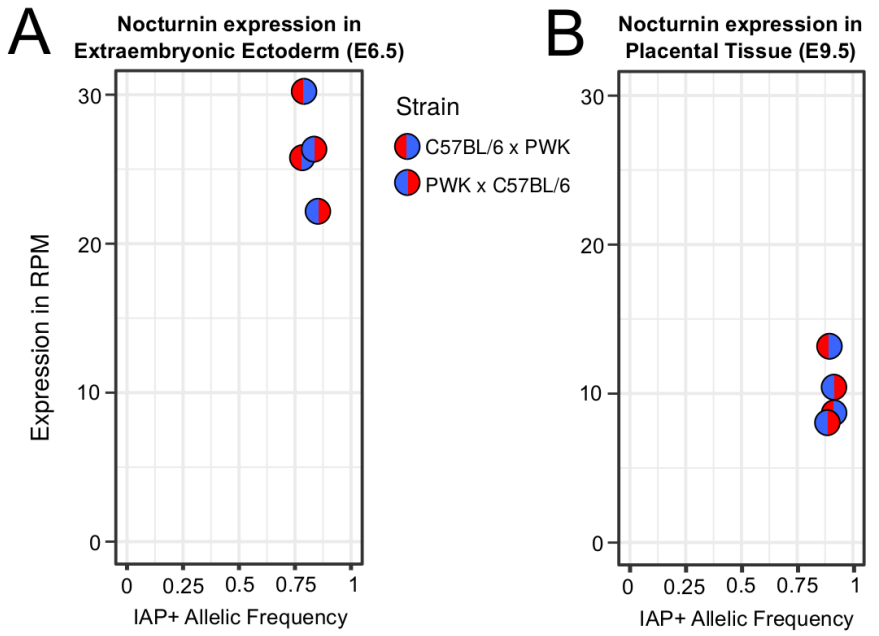


Figure 12. **Nocturnin expression in extraembryonic endoderm and placental tissue shows a specific IAP-driven expression regardless of the parent of origin.** In both A) and B) we see similar expression and similar allelic frequency (IAP+ dominated) of the reciprocal cross.

In extraembryonic ectoderm at E6.5 and placental tissue at E9.5 there is predominantly IAP+ allellic expression (data from (Inoue et al. 2017) (Fig 12 A,B). This is a clear biased allelic expression in cis regardless of the parent of origin indicating that the IAP may be what triggers the expression of this gene in placenta.

Conclusions

In conclusion, I find that the IAP plays a key role in *Noct* gene regulation in embryo development. In cis, hence as a mechanism linked to the gene locus, there is only expression of the IAP+ allele in placenta and extraembryonic ectoderm. Moreover, a parent-of-origin dependent expression in embryos means that there is inheritance of the epigenetic state, suggesting that the IAP leads to the imprinting of the gene in males. The predominantly paternal expression of the *Noct* allele with the IAP insertion is consistent in RNA-seq expression data from two independent sources. Confirming this, *Noct* IAP+ expression is higher in embryos derived from two paternal pronuclei (androgenetic embryos) than from two maternal pronuclei (parthenogenetic embryos).

Bibliography

Borensztein, Maud, Laurène Syx, Katia Ancelin, Patricia Diabangouaya, Christel Picard, Tao Liu, Jun-Bin Liang, et al. 2017. "Xist-Dependent Imprinted X Inactivation and the Early Developmental Consequences of Its Failure." *Nature Structural & Molecular Biology* 24 (3): 226–33.

Chuong, Edward B., M. A. Karim Rumi, Michael J. Soares, and Julie C. Baker. 2013. "Endogenous Retroviruses Function as Species-Specific Enhancer Elements in the Placenta." *Nature Genetics* 45 (3): 325–29.

Deng, Qiaolin, Daniel Ramsköld, Björn Reinius, and Rickard Sandberg. 2014. "Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells." *Science* 343 (6167): 193–96.

Inoue, Azusa, Lan Jiang, Falong Lu, Tsukasa Suzuki, and Yi Zhang. 2017. "Maternal H3K27me3 Controls DNA Methylation-Independent Imprinting." *Nature* 547 (7664): 419–24.

Keane, Thomas M., Leo Goodstadt, Petr Danecek, Michael A. White, Kim Wong, Binnaz Yalcin, Andreas Heger, et al. 2011. "Mouse Genomic Variation and Its Effect on Phenotypes and Gene Regulation." *Nature* 477 (7364): 289–94.

Ke, Xiayi, N. Simon Thomas, David O. Robinson, and Andrew Collins. 2002. "The Distinguishing Sequence Characteristics of Mouse Imprinted

Genes." *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 13 (11): 639–45.

Kim, Daehwan, Geo Pertea, Cole Trapnell, Harold Pimentel, Ryan Kelley, and Steven L. Salzberg. 2013. "TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions." *Genome Biology* 14 (4): R36.

Krueger, Felix, and Simon R. Andrews. 2016. "SNPsplit: Allele-Specific Splitting of Alignments between Genomes with Known SNP Genotypes." *F1000Research* 5 (June): 1479.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. "featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features." *Bioinformatics* 30 (7): 923–30.

Li, Jingfeng, Keiko Akagi, Yongjun Hu, Anna L. Trivett, Christopher J. W. Hlynialuk, Deborah A. Swing, Natalia Volfovsky, et al. 2012. "Mouse Endogenous Retroviruses Can Trigger Premature Transcriptional Termination at a Distance." *Genome Research* 22 (5): 870–84.

Luedi, Philippe P., Alexander J. Hartemink, and Randy L. Jirtle. 2005. "Genome-Wide Prediction of Imprinted Murine Genes." *Genome Research* 15 (6): 875–84.

Nellåker, Christoffer, Thomas M. Keane, Binnaz Yalcin, Kim Wong, Avigail Agam, T. Grant Belgard, Jonathan Flint, David J. Adams, Wayne N. Frankel, and Chris P. Ponting. 2012. "The Genomic Landscape Shaped by Selection on Transposable Elements across 18 Mouse Strains." *Genome Biology* 13 (6): R45.

Watanabe, Toshiaki, Shin-Ichi Tomizawa, Kohzoh Mitsuya, Yasushi Totoki, Yasuhiro Yamamoto, Satomi Kuramochi-Miyagawa, Naoko Iida, et al. 2011. "Role for piRNAs and Noncoding RNA in de Novo DNA Methylation of the Imprinted Mouse Rasgrf1 Locus." *Science* 332 (6031): 848–52.

Xie, Wei, Cathy L. Barr, Audrey Kim, Feng Yue, Ah Young Lee, James Eubanks, Emma L. Dempster, and Bing Ren. 2012. "Base-Resolution Analyses of Sequence and Parent-of-Origin Dependent DNA Methylation in the Mouse Genome." *Cell* 148 (4): 816–31.

Chapter 3: Paternal diet defines offspring chromatin state and intergenerational obesity

In this work, we describe the molecular basis behind a model of InterGenerational Metabolic Reprogramming (IGMR) based on epigenetic inheritance in *D. melanogaster*. The tasks included in this thesis comprise the bioinformatic analyses presented in the published article. By analyzing next-generation sequencing data I describe gene expression changes that are a consequence of parental exposure a dietary intervention and correlate them with chromatin architecture.

To describe changes in gene regulation given by parental diet I analyze embryos sired from flies that were fed either a medium-sugar or a high-sugar diet (3 replicates of each diet). I identify changes in gene expression consistent with a de-silencing of heterochromatin. Using Gene Set Enrichment Analyses (GSEA) I identify overexpressed metabolic pathways related to energetic metabolism and lipid storage, consistent with the increased adiposity of these flies. I also identify a downregulation of pathways related to chromatin silencing.

To test chromatin association with the metabolic response I intersect our gene expression data with the genomic segmentation (Filion et al. 2010) into 5 chromatin types classified as colors. I find an enrichment of overexpressed genes in the “black” lamin- and H1- associated heterochromatin and “blue” polycomb-associated chromatin. I also observe a depletion for “yellow” chromatin, associated with highly expressed, house-keeping genes.

For a more in depth analysis of the mechanism involved in this dysregulation I compare our embryonic IGMR RNA-seq data with gene expression datasets of H3K9me3 mutant [Su(var) 3-9, SetDB1, and HP1] embryos (Lundberg et al., 2013) and Polycomb- and Pho-RNAi knockdown embryos (Goodliffe et al., 2007). Our IGMR model showed a 70% overlap of upregulated transcripts with Su(var)3-9, SetDB1 and HP1 mutants and Polycomb--insufficient animals. Moreover, similar metabolic pathways appeared dysregulated in the mutants’ datasets when compared to wild-type. This data indicates that intergenerational metabolic responses are chromatin-state

dependent and that Polycomb- and H3K9me3- chromatin regulators are required for paternal diet induced intergenerational obesity.

Next, I analyze RNA-seq from sperm of flies fed either of the two diets (2 replicates of each diet). Consistent with our previous results, I identify dysregulation of genes embedded in “black” chromatin, suggesting that the chromatin-dependent signatures of IGMR are forecast in the paternal germline.

To further validate this hypothesis, I use modEncode H3K9me3 and H3K27me3 profiles from the same stage embryos (16-20h) and from an earlier time point (12--16h) (Negre et al. 2011). This allowed us to profile the gain of each histone modification during this highly dynamic period in the embryo: the bodies of our IGMR genes are unmarked in 12--16hr embryos but exhibit strong H3K9me3 and H3K27me3 just 4 hours later. Thus, genes undergoing highly dynamic H3K9me3 and H3K27me3 dependent silencing are specifically targeted for IGMR de-repression.

Last, I scan the context of insulator occupancy (Negre et al. 2010) of our IGMR- dysregulated genes. I find that these genes are on average farther than expected from class I insulators (CTCF, CP190, and BEAF-associated) and somewhat closer to class II (SuHw-associated) insulators.

Altogether, these findings support a model of reprogramming based on a redistribution of heterochromatin in the sperm and zygote. We conclude that IGMR is characterized by H3K9me3- and Polycomb-dependent dysregulation.

Paternal sugar alters offspring heterochromatin

Paternal IGMR appeared phenotypically "silent" through the complexities of development. We therefore hypothesized that the phenotype was encoded in chromatin. Position-effect varie- gation (PEV) is a genetic phenomenon that has been used as a quantitative readout of locus-specific chromatin state silencing in vivo. The most common PEV reporters in Drosophila reflect chromatin desilencing through increased expression of a red- eye-pigment-coding reporter gene. Screening a library of PEV strains, Phalke and colleagues recently defined at least five functionally distinct chromatin silencing subtypes in the living fly (Phalke et al., 2009). Intriguingly, when testing w^{m4h}, a reporter for peri-centric heterochromatin on ChrX, we observed a reproducible U-shaped intergenerational eye color phenotype. In support of a mechanistic link between the IGMR obesity and PEV results, triglyceride accumulation and eye color correlated positively; redder-eyed flies were more obese. No correlation was observed in the remaining four strains. These data show that acute paternal diet targets select chromatin subtypes in offspring.

At this point, we focused on medium- versus high-sugar IGMR and tested whether IGMR affected all or only select individuals in the population. Measuring pigment from single wm4h fly heads as a direct readout of the IGMR response, we observed that paternal IGMR red-shifted the entire distribution (Fig. 13A and 13B). Thus, high paternal sugar induces w^{m4h} desilencing population-wide, indicating that each paternal gamete carries an equivalent intergenerational signal.

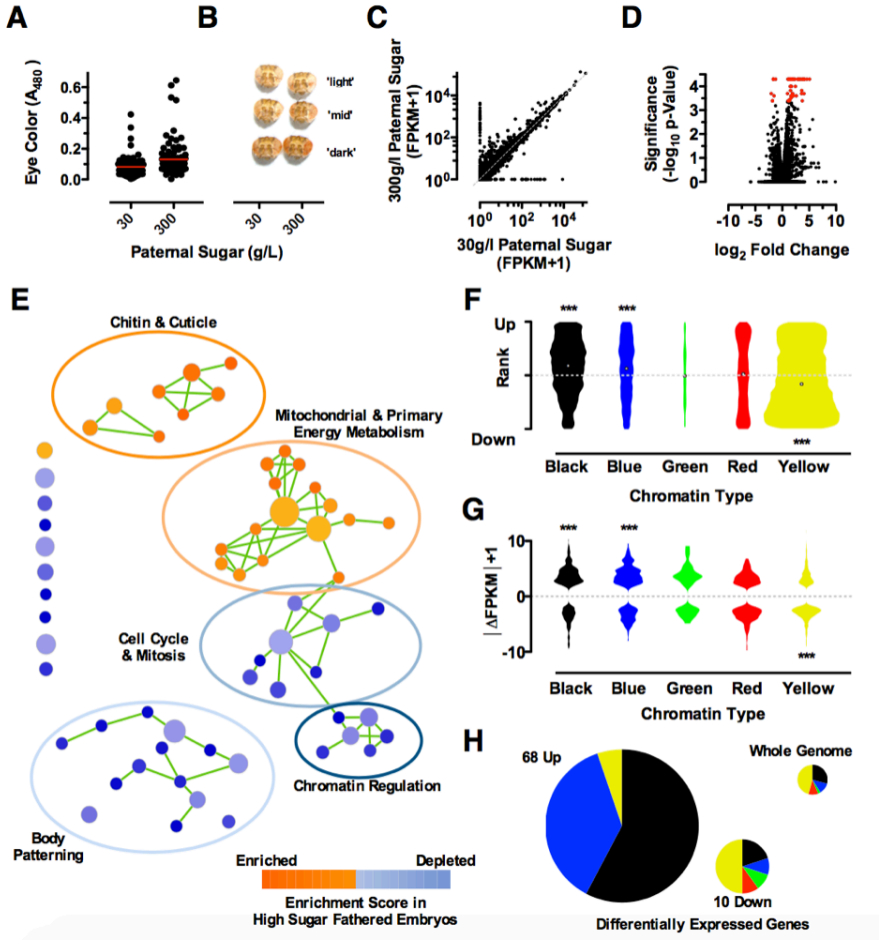


Figure 13. Paternal IGMR alters select chromatin states in offspring. A) Interindividual variation of eye color of w^{m4h} flies from fathers fed high (300 g/l) or medium (30 g/l) sugar food. B) Representative heads from wm4h offspring. Shown are heads representative of light, medium, and dark red eyes of each respective cohort. (C–H) RNA-sequencing results of medium (30 g/l) and high-sugar (300 g/l) sired stage 17 embryos. C) FPKM plot. D) Volcano plot. E) Cytoscape enrichment map (p cutoff: 0.005, FDR Q-value cutoff: 0.025, overlap cutoff: 0.2) of gene set enrichment analysis (GSEA). (Orange) Gene sets enriched; (blue) gene sets depleted, in high-sugar IGMR. Color intensity reflects degree of enrichment. Major clusters are circled. (F and G) (F) Rank and (G) absolute IGMR expression changes. Genes are allocated to one of five chromatin states (colors) according to their TSS (Filion et al., 2010). Plotted are (F) ranks for all genes and (G) absolute expression changes of the top 1,000 IGMR up and downregulated genes. H) Chromatin color annotation of all significantly up- and downregulated IGMR genes.

We performed rRNA-depleted RNA-sequencing of hand-picked stage 17 embryo F1 offspring from medium- and high-sugar challenged fathers (Pearson corr. = 0.97, ~15 million reads/sample; Fig. 13C). In support of a selective chromatin state desilencing mechanism, gene expression broadly increased, with many more up- than downregulated transcripts. Sixty-eight protein-coding genes were significantly upregulated in high-sugar sired embryos (mean ΔFPKM = 54.9) and only ten downregulated (mean ΔFPKM = 7.0; Fig. 13D and Table S1). Of note, upregulated transcripts tended to be genes highly expressed during late embryo and early larval stages, including 27 (40%) related to biogenesis of the sugar-based cuticle. Of the remaining 42 genes, 30 were of unknown function, 5 had peptidase activity, and interestingly, 4 were metabolic genes, including fatty acyl-CoA reductase and fatty acid elongase.

Analysis using gene set enrichment analysis (GSEA) revealed two clearly upregulated clusters containing chitin and cuticle constituent and mitochondrial and primary energy metabolism pathways (Fig. 13E). Included and consistent with the heightened adiposity of IGMR, pathways for lipid particle, the electron transport chain complexes I, IV, and V, glycolysis, TCA cycle, and fatty acid metabolism were all upregulated. These changes are consistent with energetics of enhanced lipid storage. Three downregulated clusters were also detected, including cell cycle and mitosis, body patterning, and intriguingly, a cluster of chromatin regulation pathways. Consistent with sensitivity of the pericentric w^{m4h} reporter to IGMR chromosome, “centromeric region” was ranked second in the chromatin cluster and “chromatin silencing” ranked third. Examination of genes annotated as PEV suppressing, also known as Su(var)’s, revealed a concerted ~10%–20% downregulation, including members of most well-documented silencing pathways. Thus, the IGMR embryo is characterized by gene expression favoring primary energy metabolism over chromatin control.

We next compared our data with chromatin mapping data sets from the community. Filion et al. used DamID to annotate five ma-or chromatin types, three repressive (black, blue, and green) and two active (red and yellow) (Filion et al., 2010). When intersecting our IGMR embryo data with their chromatin state maps, strong

enrichment was observed in high-sugar sired embryos for genes embedded in “black” lamin/H1-associated heterochromatin and “blue” polycomb-associated chromatin, and relative depletion was observed for those annotated as “yellow”, or housekeeping-type chromatin (Fig. 13F and 13G). These findings were verified using rank-order (Fig. 13F) and differential expression analyses (Fig. 13G). No global effect was observed on “red” or “green” chromatin embedded genes. Consistent with these global indications of chromatin state dependency, the 68 significantly upregulated genes were almost exclusively found in “black” or “blue” chromatin while the 10 significantly downregulated transcripts were randomly distributed (Fig. 13H). These data identify high paternal sugar as a chromatin-state-selective physiological Su(var) and identify IGMR as chromatin state dependent.

Polycomb and core heterochromatin machinery mediate paternal IGMR

To genetically validate chromatin state regulation as a mechanistic underpinning of our model, we began systematically testing IGMR potential in mutants known to modify w^{m4h} variegation. We started with Su(var)3-906, a homozygous dominant suppressor allele of the H3K9 histone methyltransferase Su(var)3-9.

Medium- and high-sugar-challenged Su(var)$3\text{-}9^{06}$ fathers were mated with standardized w1118 females, and the resulting het- erozygote offspring were monitored for adiposity (Fig. 14A). Whereas w^{1118} animals reproducibly exhibited a ~10%–15% increase in adiposity upon high-sugar IGMR, F1 adult male offspring of Su(var)$3\text{-}9^{06}$ fathers showed no intergenerational obesity response (Fig. 14A). This provides genetic evidence that Su(var)3-9 is required for IGMR.

We also tested a second H3K9 methyltransferase, SetDB1. As heterozygotes, $SetDB1^{1473}$ fathers gave both wild-type and mutant offspring. Intriguingly, both mutant (Fig. 14B, red) and wild-type $SetDB1^{1473}$ fathered offspring (Fig. 14B, black) completely failed to mount an IGMR obesity response (Fig. 14B). Drosophila sperm develop as a syncytium, and therefore both mutant and wild-type sperm in such a cross will share a $SetDB1^{1473}$ mutant cytosolic

compartment for most of their development. These findings therefore indicate that SetDB1 in the male germline is necessary for proper IGMR.

H4K20me3 deposition follows H3K9me3 in the establishment of heterochromatin (Schotta et al., 2004). We therefore also tested Su(var)4-20^{SP}, a mutant for the H4K20 methyltransferase Su(var)4-20. As Su(var)4-20^{SP} is on Chr X, all male offspring from our crosses are wild-type. Again though, wild-type offspring will reflect the mutant heterozygosity of spermatogenesis. Su(var)4- 20^{SP} fathers failed to transmit paternal IGMR to the F1 (Fig. 14C). Thus, uncompromised expression of Su(var)3-9, SetDB1, and Su(var)4-20 are absolutely required for IGMR. Of note, not all wm4h suppressor alleles were IGMR incompetent. Su(var)3-1^{04} and Su(var)3-3^{12}, also known as Jil1 kinase and dLSD1, respectively, generated completely normal IGMR obesity responses (Fig. 14D), thus indicating that IGMR is not directly linked to the wm4h insertion locus itself. These findings identify one of the first gene networks known to be absolutely required for proper intergenerational metabolic reprogramming. Given the observed derepression in blue embedded genes (polycomb-associated; Fig. 13), we tested IGMR potential in polycomb and trithorax group mutants. We found that, although Ash1^{22} mutants were fully IGMR competent, Enhancer of zeste, E(z)^{63} and Polycomb, Pc^{3} mutant males completely failed to elicit a response in the next generation (Fig. 14E). Thus, polycomb- and H3K9me3-centric chromatin regulators are absolutely required for paternal diet-induced intergenerational obesity.

The IGMR program is chromatin encoded

To corroborate these findings, we compared our embryonic IGMR RNA-seq data with profiles from H3K9me3- and poly- comb-insufficient mutants. We examined profiles from Su(var) 3-9$^{06/evo}$, SetDB1$^{10.1}$, and HP1$^{04/05}$ mutant first-instar larvae (Lundberg et al., 2013) and Pc- and Pho-RNAi knockdown em- bryos (Goodliffe et al., 2007).

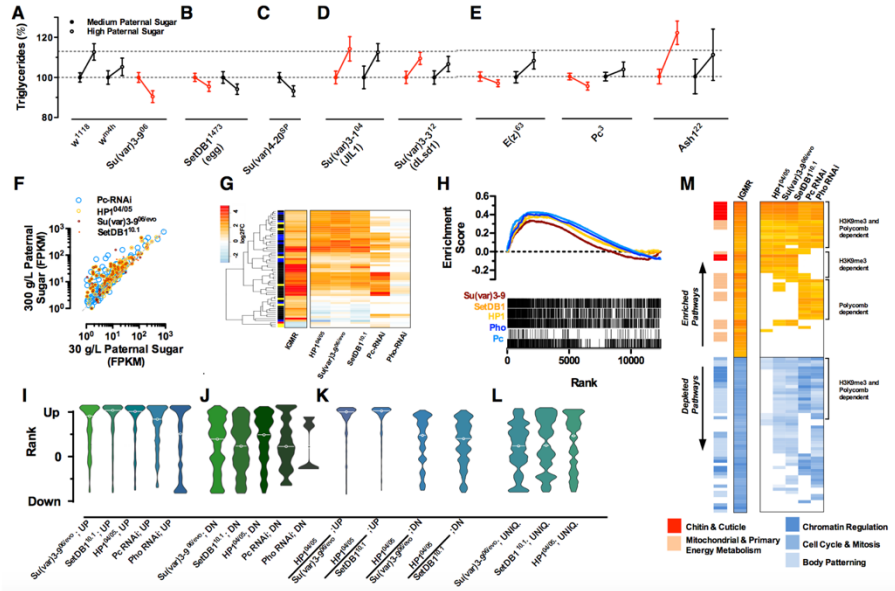


Figure 14. A Su(var)/PcG axis essential for paternal IGMR. (A–E) Adiposity of offspring (triglycerides/weight) of mutant fathers challenged with medium (30 g/l; closed circles) or high sugar (300 g/l; open circles). Gray dashed line indicates normal w^{1118} IGMR response. IGMR adiposity responses are shown for offspring of (A) w^{1118}, w^{m4h}, and Su(var)3-9^{06}, (B) SetDB1^{1473}, (C) Su(var)4-20^{SP}, (D) Su(var)3-1^{04}, Su(var)3-3^{12}, (E) E(z)^{63}, Pc^{3}, and Ash1^{22} mutant (red) and wild-type (black) offspring. Results are mean ± SEM of n= 3–8 experiments each with multiple replicates. F) FPKM values of RNAseq data from medium and high-sugar-fathered embryos. 200 most upregulated genes from HP1 (yellow), Su(var)3-9 (cayenne), and SetDB1 (orange) mutant first-instar larvae from Lundberg et al. and Pc-RNAi (blue) experiment from Goodliffe et al. G) Heatmap of expression changes of significantly changed genes in our paternal IGMR offspring embryo data set and in the Lundberg et al. HP1, Su(var)3-9, and SetDB1 mutants and the Goodliffe et al. Pc-, Pho-RNAi data sets. H) Enrichment plot for gene sets upregulated in HP$1^{04/05}$ (yellow), Su(var)3-$9^{06/evo}$ (cayenne), and SetDB$1^{10.1}$ (orange) mutants and Pc- (light blue) and Pho-RNAi (dark blue) in our stage 17 paternal IGMR offspring embryos. (I–L) Violin plots of expression change distributions relative to all genes of stage 17 paternal IGMR offspring embryos for gene sets from HP$1^{04/05}$, Su(var)3-$9^{06/evo}$, and SetDB$1^{10.1}$ mutants (Lundberg et al.) and from Pc and Pho-RNAi embryos (Goodliffe et al.). IGMR relative rank is plotted for all available of the (I) 200 genes most upregulated and (J) 200 genes most downregulated by each mutant / RNAi line; (K) intersects of the 200 most up- or downregulated genes of the indicated pairs of mutants; and (L) genes in the 200 most upregulated gene sets unique to each respective mutant. M) Heatmap comparison of GSEA results from mutant and IGMR data. Plotted are the 50 most up- and downregulated pathways from paternal IGMR and respective scores from the mutant data sets. Colored bars left of the heatmap indicate clusters in Figure 13E.

Intriguingly, ~70% overlap was observed between our significantly dysregulated IGMR genes and those responsive to H3K9-centric or polycomb insufficiency (Fig. 14F and 14G). The converse was equally true; each of the top 200 Su(var)3-9$^{06/evo}$, SetDB1$^{10.1}$, HP1$^{04/05}$, Pho-RNAi, and Pc-RNAi dysregulated gene sets showed strong enrichment in our high-sugar-sired F1 embryos (Fig. 14H). Subgrouping confirmed specificity of these signals. First, transcripts upre- gulated by Su(var)3-9$^{06/evo}$, SetDB1$^{10.1}$, HP1$^{04/05}$, Pho, and Pc insufficiency (likely direct targets) showed clear coordinate in- creases in expression (Fig. 14I) compared to apparently randomly distributed signals for transcripts downregulated by mutation (Fig. 14J). Transcripts upregulated by both HP1$^{04/05}$ and either Su(var)3-9$^{06/evo}$ or SetDB1$^{10.1}$ (Fig. 14K) showed much stronger signatures than transcripts significantly regulated by any one Su(var) mutant alone (Fig. 14L). Thus, paternal IGMR mimics H3K9me3- and polycomb-dependent transcriptional dysregulation.

To test whether these signatures might directly contribute to metabolic reprogramming, we performed GSEA analysis of the Lundberg et al. (2013) and Goodliffe et al. (2007) data sets. Coordinate overlapping enrichment signatures were observed for key pathways of all five major IGMR clusters (Fig. 14M), including most chromatin and primary energy modules. Of note, the most significantly enriched pathways in our data set were those regulated by both silencing systems together (Fig. 14M). Thus, IGMR is characterized by H3K9me3-/PcG- dependent dysregulation.

Sperm and Zygote Chromatin Plasticity Define IGMR

To gain further insight into IGMR transmission, we performed RNA sequencing from manually dissected and purified mature sperm of high- and medium-sugar-fed w^{1118} males.

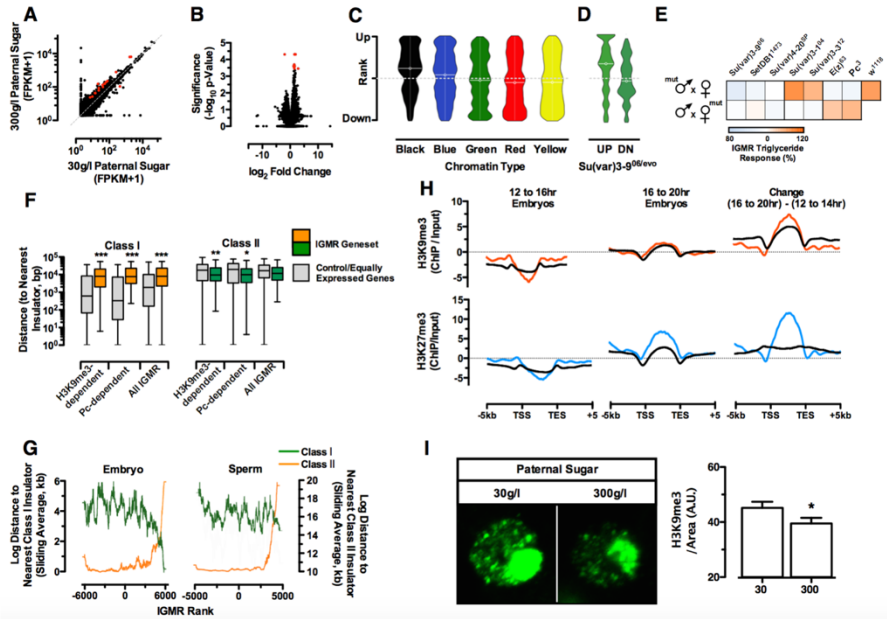


Figure 15. IGMR signatures are forecast in the P0 germline, and IGMR changed genes show K9/K27me3 dynamic context. (A–D) RNA-sequencing results of sperm from medium- (30 g/l) and high-sugar (300 g/l) fed fathers; significantly changed genes are depicted in red. A) FPKM plot. B) Volcano plot. C and D) IGMR expression changes in sperm of high- relative to medium-sugar-fed fathers (C) for the five chromatin colors according to Filion et al. and (D) for 200 most up- or downregulated genes from Su(var)3-$9^{06/evo}$ mutants from Lundberg et al. E) Relative adiposity of male offspring (triglycerides/weight) (top row) from crosses of mutant fathers challenged with medium- or high-sugar diet, with w^{1118} mothers and (bottom row) of crosses of w^{1118} fathers with mutant mothers. The normal w^{1118} IGMR response is also shown (top row). Results are mean ± SEM of n= 3–8 experiments each with multiple replicates. F) ChIP/input signal from modENCODE data sets for leading-edge H3K9me3- and polycomb-dependent genes (red in top panels, blue in bottom panels) in our IGMR offspring embryo RNA-seq. H3K9me3 (top) and H3K27me3 (bottom) enrichment of 12- to 16-hr-old and 16- to 20-hr-old embryos (left) and the difference between the two stages (right). Black lines present the average for all genes. G) Box plots of distance to nearest class I and class II insulators. Shown are distances for leading-edge Su(var) and PcG upregulated genes in our IGMR offspring embryo RNA-seq. Grey boxes represent a control set of equally expressed genes. The boxes indicate the first and third quartiles, and the central line indicates the median. Whiskers extend to the most extreme data point, which is no more than 1.5 times the quartile range. H) Distance to nearest class I (orange) and class II (green) insulator plotted according to ranked expression change from IGMR RNA-seq results (high versus medium sugar). Values are sliding window averages of 500 genes. I) H3K9me3 staining of fat body cell nuclei from offspring of medium (30 g/l) and high-sugar (300 g/l) fed fathers. Results are mean ±SEM of n = 7 experiments, each with multiple replicates. Quantification of fat body cell nucleus H3K9me3 staining.

Intriguingly, we again observed clear evidence of (1) broad transcriptional derepression in sperm of high-sugar-fed males (Fig. 15A and 15B), (2) selective upregulation of black chro- matin-embedded genes (Fig. 15C), and (3) upregulation of Su(var)3-9$^{06/evo}$-sensitive genes (Fig. 15D). These data indicate that transcriptional dysregulation in mature IGMR sperm is also chromatin state defined. In contrast to the embryo data, blue and yellow embedded genes appeared largely unaffected in the sperm transcriptome. Thus, chromatin-dependent signa- tures of IGMR are forecast in the P0 paternal germline.

Dysregulation of black embedded genes in both the sperm and zygote suggested potentially overlapping mechanisms for generation of the intergenerational signal in the germline and for hardwiring the IGMR phenotype in the offspring. To probe this idea genetically, we compared the effect of maternal versus paternal mutant allele contribution on IGMR. As described above, offspring of Su(var) and Polycomb mutant fathers were incapable of mounting an IGMR response (Fig. 14A–14E and 15E, top row). In crosses in which Su(var)3-9$^{06/evo}$, SetDB1$^{10.1}$, or Su(var)4-20^{SP} mutations were contributed by the oocyte, IGMR-competent wild-type sperm were no longer able to evoke an intergenerational response (Fig. 15E, bottom row). Su(var)3-1^{04} and Su(var)3-3^{12} mutants, unremarkable in the male germline, completely abrogated the response when contributed maternally (Fig. 15E, bottom row). In contrast, oocytes contributing Pc^{3} and E(z)^{63} mutations, whose constitutive heterochromatin would not be predicted to be directly perturbed, mounted completely normal IGMR responses. Collectively, these data support a model in which IGMR results from and requires a permissive range of heterochromatin plasticity in the zygote.

To validate the idea, we intersected our embryo RNA-seq data with modENCODE H3K9me3 and H3K27me3 ChIP-seq profiles from same-stage embryos (16–20 hr) and from those isolated one time point earlier in development (12–16 hr), enabling us to gauge the dynamics of K9me3/K27me3 gain and loss (Negre et al., 2011). We made several observations. First, IGMR-dysregulated genes represented a class undergoing highly dynamic H3K9 and H3K27 trimethylation (Fig. 15F). This was true for our significantly changed IGMR genes, as well as the

leading edge H3K9me3- and polycomb-dependent IGMR gene sets from Fig. 14I (Fig. 15F and data not shown). The bodies of these genes in particular were unmarked in 12–16 hr embryos and exhibit strong H3K9me3 and H3K27me3 just 4 hr later. Importantly, we observed the same signature when analyzing leading-edge genes of metabolic pathways upregulated in our obese IGMR phenotype. Thus, genes undergoing highly dynamic H3K9me3- and H3K27me3-dependent silencing are specifically targeted for IGMR derepression.

Because repressive marks correlate with the higher-order chromatin structure and cis-regulatory domain organization, we also examined our gene sets in the context of insulator occupancy (Negre et al., 2010). Analysis revealed that all three IGMR- dysregulated gene sets were on average far from class I (CTCF, CP190, and BEAF-associated) and were somewhat closer to class II (SuHw-associated) insulators (Fig. 15G). These signatures were specific when compared to similarly expressed genes or to the entire transcriptome (Fig. 15H, left). Intriguingly, the same signature was again evident in our most up- and downregulated sperm transcripts (Fig. 15H, right). Thus, IGMR impacts spatially and chromatin-context-defined transcriptional units in fathers and in offspring.

Collectively, our data suggest that IGMR results from global alterations in chromatin state integrity within a permissive window, where obesity susceptibility results from reduced stage-specific epigenetic regulation of H3K27me3- and H3K9me3-defined domains. Our observations of wm4h eye color desilencing (Fig. 13) and reductions in H3K9me3 immunofluorescence in adult IGMR offspring fat bodies (Fig. 15I) indicate that this chromatin state reprogramming is stable lifelong in the offspring.

Chapter 4: Transgenerational transmission of environmental information in C. elegans

Here, we describe a model of transgenerational epigenetic inheritance based on elevated expression of a multicopy transgene in *C. elegans*. The transgene is normally silenced by SET25-dependent H3K9me3, which is inhibited after exposure to 25C. After high-temperature exposure the transgene remains upregulated for many generations even when brought back to 20C. Notably, we find endogenous genes that behave in a similar way.

The tasks included in this thesis comprise the bioinformatic analyses presented in the published article. By analyzing next-generation sequencing data I describe gene and transposon expression changes that are a consequence of parental growth at high temperature and suggest a role for heterochromatin in transgenerational epigenetic inheritance.

I analyze total RNA sequencing data from worms grown at 25C and at 20C (P0) to identify genes that show elevated expression after environmental growth at high temperature. I analyze total RNA sequencing data from the 3rd generation (F3) progeny of these worms, all maintained at 20C for the 3 generations, and I identify several pseudogenes and DNA transposons that maintain memory of elevated expression after environmental growth at high temperature.

To prove the hypothesis that genes that maintain memory of upregulation in the 3rd generation are also normally repressed by H3K9me3 I analyze total RNA sequencing data from *set25* mutant worms. I identify the same pseudogenes and DNA transposons that are upregulated after exposure to high temperature in *set25* mutants. The significant overlap, correlation and prediction value between memory of expression after 3 generations and lack of H3K9me3 suggest a role for this histone modification in maintaining transgenerational epigenetic information of gene expression.

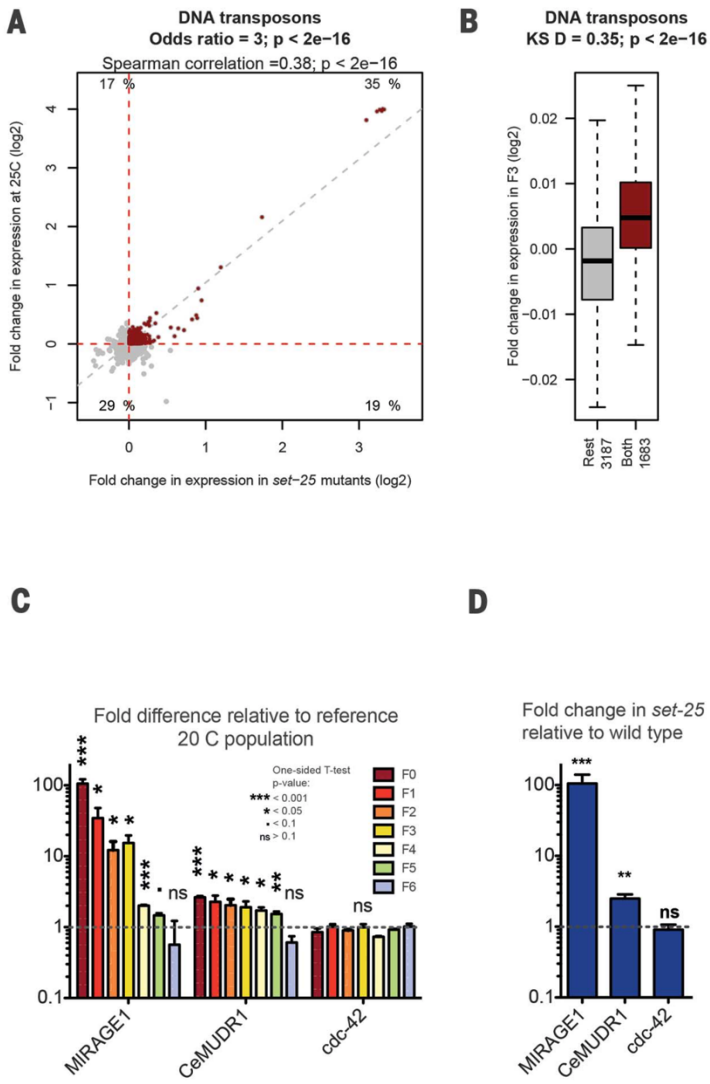


Figure 16. Epigenetic expression memory of endogenous loci repressed by SET-25. A) DNA transposon expression change in set-25 mutants and at high temperature. Odds ratio quantifies the overlap (red loci, "both") between log2 fold change (FC) > 0. B) FC expression three generations after a reduction in temperature from 25° to 20°C. Kolmogorov–Smirnov (KS) test statistic and P value are shown. See figs. S11 to S14 and table S4 for other repeats, protein coding genes, and analysis methods. C) Expression of two DNA transposons at 25°C (F0) and for six generations after decreasing the temperature to 20°C determined by quantitative PCR (table S3). cdc-42 is a housekeeping gene as control. Expression is relative to animals grown at 20°C in parallel. D) Expression of the same DNA transposons is increased in set-25 mutants. **P < 0.01; ***P < 0.001; ns, not significant.

Our results suggest a simple model for how the transgene array shows memory of high-temperature exposure that endures for many generations. High temperature inhibits SET-25-mediated repression in the germ line, causing loss of H3K9me3 from the array. This derepressed chromatin is transmitted to subsequent generations, resulting in increased expression when transcription initiates in somatic lineages. Over multiple generations of growth at low temperature, repression is gradually restored by heterochromatin remodeling in each germline cycle. This is consistent with previously reported gradual quantitative intergenerational changes in H3K9me3 following a temperature change at some loci (14).

We tested whether this model predicts the behavior of endogenous loci in the genome by sequencing RNA from set-25 mutants and WT animals at 20° and 25°C and from WT animals three generations after a change from 25° to 20°C. For protein-coding genes, derepression in set-25 mutants provided weak prediction of increased expression at high temperature, consistent with a larger contribution from other regulators, such as specific transcription factors. Derepression in set-25 mutants was, however, a better predictor of increased expression at high temperature for multiple classes of repetitive elements and also for pseudogenes (Fig. 16A), consistent with impaired SET-25 activity's making an important contribution to the increased expression of many loci at high temperature. Moreover, the increased expression of loci repressed by SET-25 with increased expression at high temperature was still detectable three generations after a return to low temperature (Fig. 16, A and B). Quantifying the expression of two DNA transposons by quantitative real-time polymerase chain reaction (PCR) in independent samples confirmed that their expression remained elevated for five generations after a return to 20°C (Fig. 16C). Their expression was also confirmed as SET-25-dependent (Fig. 16D).

These results support the mechanistic model: At high temperature, SET-25 pathway activity is reduced, resulting in the derepression of many loci in the genome. After a return to low temperature, SET-25 activity is restored, but it takes multiple generations for repression to be reestablished. Expression from SET-25-repressed repeats transmits information about a prior environmental exposure in this species.

Chapter 5: Impaired DNA replication derepresses chromatin and generates a transgenerationally inherited epigenetic memory

In this chapter I describe a model of transgenerational epigenetic inheritance based on upregulation of a multicopy transgene in *C. elegans*. The transgene is normally silenced by H3K27me3 and H3K9me3. These histone modifications are lost during impaired DNA replication and as a consequence the transgene is overexpressed. The modified chromatin is transmitted for many generations before fully resetting. Many endogenous genes repressed by heterochromatin also see their expression highly increased after impaired replication during embryonic development.

The tasks included in this thesis comprise the bioinformatic analyses presented in the published article. By analyzing next-generation sequencing data I describe gene expression changes consequence of heterochromatin loss after impaired DNA replication. I use chromatin datasets from modEncode to map the gene expression changes to chromatin subtypes and identify an enrichment of upregulation in the 4 heterochromatin subtypes. Specifically, I intersect the coordinates of all genes with the coordinates of the chromatin segmentation for the same stage in modEncode data. I use the chromatin state that covers the largest fraction of the gene. However, using the chromatin state that spans the gene promoter doesn't change the outcome.

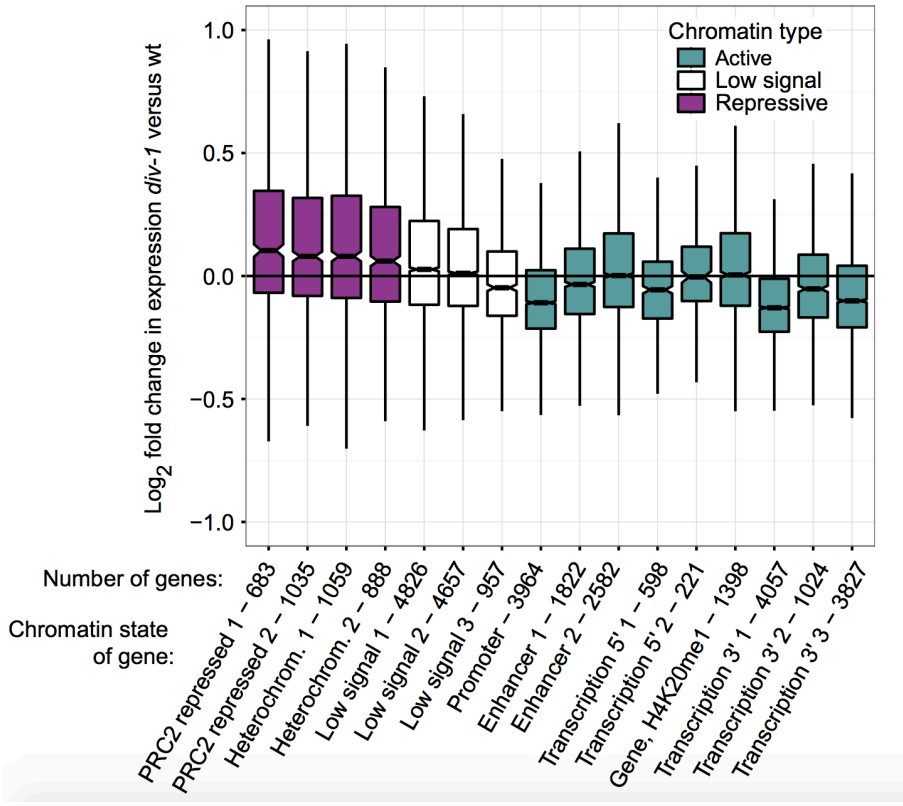


Figure 17. Impaired DNA replication globally derepresses chromatin. Fold change in expression of genes mapping to different modENCODE chromatin states between div-1 and wt L1 larvae. The number of genes assigned to each state is indicated.

To test whether loci other than the transgene array also have altered expression when replication is impaired, we sequenced RNA from WT and *div-1* L1 stage larvae. Consistent with the response of the transgene array, many more genes had increased compared to decreased expression in the div-1 mutants [493 up-regulated genes versus 9 down-regulated genes at a false discovery rate (FDR) of <0.05].

To relate changes in expression to the normal chromatin state of each gene, we used data from the modENCODE consortium (15). Consistent with the response of the array, this revealed widespread up-regulation of genes with normally repressed chromatin states (Fig. 17). This derepression was observed for genes normally characterized

by states defined by either high H3K9me2/3 or high H3K27me3 (Fig. 17). In contrast, genes without repressive chromatin states (15) in WT animals were not up-regulated as a group in *div-1* mutants (Fig. 17). Together, these results show that impaired DNA replication during early development has a major impact on chromatin and gene expression, globally reducing the levels of repressive histone modifications and causing widespread up-regulation of heterochromatic genes in the resulting animals.

Discussion

Repetitive elements are present in the vast proportion of the genome that does not belong to protein coding genes and that is usually considered 'junk'. TE have to be repressed to maintain genome integrity. Cells usually achieve TE silencing by DNA methylation and histone modifications among other epigenetic mechanisms, which we know can be affected by environmental cues contributing to phenotypic diversity.

In the first chapter of this thesis I analyse inter-individual variation in piRNA expression, aiming to associate piRNA expression with genetic polymorphisms and ultimately gain insight into mammalian piRNA biogenesis and evolution. The data and experimental results are generated in collaboration with J. Jimenez-Chillaron lab (from Fundació Sant Joan de Déu) and Sonia Forcales lab (from the IGTP).

To do so, I analyze small RNA sequencing data from testes of different mouse individuals from an outbred mouse strain, and I compare expression of each piRNA cluster between each individual. Here, I find one piRNA cluster with bimodal expression: it is expressed in some animals and silent in some others. Interestingly, this piRNA cluster overlaps a protein-coding gene, Noct, that is expressed in germline tissue and with a known polymorphic IAP insertion.

I associate the expression of the piRNA cluster with animals that have the IAP element. To substantiate our findings I analyse testis small RNA data from inbred mouse strains that are either IAP- (FVB, C3H and 129S mouse strains) or IAP+ (C57BL/6 and NOD mouse strains), and in all cases the presence of the IAP matches the expression of the piRNA cluster. According to our hypothesis, the IAP casts the transcript into the piRNA biogenesis pathway. Specifically, it is the IAP-containing intron that is recognized as a repetitive element and cleaved into piRNAs. Thus, I present a novel function of IAP elements: they can switch a protein coding gene into a piRNA cluster. I describe a model where a genomic integration of a naturally occurring IAP element directs an intron into piRNA processing.

Importantly, other intronic IAP insertions in germline-expressed protein-coding genes are also associated with piRNA production from intronic transcripts. The massive expansion of IAP endogenous retroviruses in laboratory mouse strains has thus triggered the birth of new piRNA producing loci, some of which are not yet fixed in the species. Our work reveals how genetic variants shape the piRNA repertoire in a mammalian species.

Last, I identify a nuclear export factor, NXF1, required for the expression of some IAP-containing piRNA clusters. *Mus musculus castaneus* (CAST) mice contain a natural mutant allele of the NXF1 gene that suppresses phenotypes associated with IAP insertional mutagenesis and the level of IAP-induced aberrant genic transcripts, although how this is achieved remains unclear (Concepcion et al. 2015). NXF1 is involved in piRNA production in flies (Dennis et al. 2016). I notice that some IAP-derived piRNA clusters are silent in CAST, suggesting that a fully functional NXF1 protein is required for piRNA production from IAP-associated piRNA precursors. This is the first time that NXF1 is involved in piRNA biogenesis in any vertebrate.

However, the exact mechanism by which RNA transcripts including the IAP (introns) manage to hijack NXF1 to be exported to the cytoplasm still remains a mystery. One plausible hypothesis is that an ancient export motif from the IAP is specifically recognized by NXF1.

In summary, I hypothesize that new piRNA producing loci appear from germline-expressed genes that are designated for piRNA biogenesis after an IAP insertion. This is consistent with previous observations that protein-coding transcripts producing piRNAs from their 3' UTRs frequently contain repetitive sequences (A. A. Aravin et al. 2007). piRNA-producing loci are therefore dynamic and reflect private transposable elements and other genetic variants present in the different individuals of this species. It is tempting to speculate that piRNA producing loci evolved from transposable element insertions in germline-expressed genomic loci.

The known function of piRNAs is to repress transposons. And the expression of the overlapping gene is not affected in male germline. So, what is the role of the piRNAs from a newly evolved piRNA cluster?

And more importantly, can the inheritance of Noct piRNAs lead to a differential expression of this gene in offspring? In fact, it is known that maternally inherited piRNAs confer immunity to repetitive elements in a chromatin-independent way in flies (Brennecke et al. 2008; Grentzinger et al. 2012). Moreover, inherited piRNAs can also trigger expression of more piRNAs by changing the chromatin state of piRNA clusters (Le Thomas et al. 2014).

Furthermore, IAP and other transposons positively regulate gene expression in embryos by driving histone modifications associated with transcription (Chuong et al. 2013). Moreover, ERVs can lead to a huge reduction of non-terminated transcripts when inherited from the father but not from the mother, thus depending on the parent of origin (Li et al. 2012).

In the second chapter of this thesis I show how an IAP element can regulate gene expression in embryos. Comparing allelic expression of hybrids for the Nocturnin IAP, I show that the IAP drives expression of *Noct* in placenta. This can be due to the global hypomethylation features in placenta in comparison to other embryonary tissues. ERV variation was already suggested as a potential mechanism of enhancer polymorphism in placenta. This allows for different developmental phenotypes and acts as an evolutionary mechanism of placental diversification via gene regulation (Chuong et al. 2013).

Then, I present an IAP-driven parent-of-origin dependent expression of the *Noct* gene. While the ancestral Nocturnin mouse allele is equally expressed from both the maternally and paternally inherited copy, the derived allele that carries the IAP insertion is predominantly expressed from the paternal copy at the blastocyst stage of mouse embryonic development.

Considering how recent this IAP insertion is (~100 years, Nellåker et al. 2012) it is extremely unlikely that another variant linked to the IAP insertion is responsible for the emergence of this parent-of-origin dependent expression pattern, these results strongly suggest that the IAP itself is the sequence element responsible for acquisition of imprinted expression at this mouse locus. The observation that an IAP insertion in mouse gene has led to the emergence of a novel imprinted

gene matches previous associations between ERVs and imprinting in both human and mouse (Luedi, Hartemink, and Jirtle 2005). To note, LTR transcription is known to drive species-specific DNA methylation in oocytes, specifically in promoters overlapping CpG islands, in patterns that can be heritable in blastocysts and placenta (Brind'Amour et al. 2018). Yet, the described mechanism does not explain our observations since we find that the imprinted IAP is inherited from the father and not from the mother. Our results suggest that there is an additional mechanism that links IAPs with imprinting independent of oocyte LTR-initiated DNA methylation.

Interestingly, there are examples of piRNAs contributing to imprinting and gene regulation. In the *Rasgrf1* locus, piRNAs drive methylation of the ICR leading to gene activation (Watanabe et al. 2011). In the *Igf2* locus, imprinted expression is controlled by a regulatory element downstream of the gene. When this imprinting control element (ICE) is not methylated, it is bound by CTCF which blocks the interaction of a downstream enhancer with the promoter of *Igf2* (Bell and Felsenfeld 2000). So, the unmethylated ICE causes repression of *Igf2*. When the ICE is methylated, CTCF does not bind and the downstream enhancer can activate the gene. In summary, methylation at the ICE leads to expression of *Igf2*. In this case, the methylated and expressed *Igf2* allele is also inherited from the father.

Last, it is worth mentioning that a recent study shows that deleting a natural ERV (Bogutz et al., n.d.) from upstream of two genes erases imprinting in offspring and leads to biallelic expression, proving a role for repetitive elements in imprinting and epigenetic inheritance. Moreover, it was shown that ERVs contribute to 3D chromatin architecture and drive gene expression at 2-cell stage embryos (Kruse et al., n.d.) as well as in human totipotent stem cells (Zhang et al. 2019). It will be interesting to test if lineage-specific polymorphic transposons are generally associated with emergence of parent-of-origin dependent expression in mouse and in humans. Importantly, these ERV can act as evolutionary mechanisms by regulating gene expression during early development.

In the following three chapters of the thesis I aim to describe the molecular mechanisms responsible for epigenetic inheritance. To do so, I analyse two different animal models with the goal of understanding the epigenetic basis behind the transmission of physiological environmental stimuli to the next generation(s).

The goal of the third chapter of this thesis is to identify a paradigm of TEI in *D. melanogaster* and describe the mechanisms behind the inheritance of the acquired complex traits. In order to do so, we design an experiment in collaboration with Pospisilik lab at the Max-Planck Institute in Freiburg, Germany and we find that these flies show predisposition to increased adiposity and obesity, with a significant gain in body weight, when challenged with a high sugar diet (Öst et al. 2014). We call this response intergenerational metabolic reprogramming (IGMR). Interestingly, the phenotype is transmitted after an acute stimulus of only two days. However, the transmission can be abolished after a heat-shock just one hour before mating, meaning that the sperm epigenome is plastic and rapidly affected. We find that fruit flies mutants for genes that are key factors of heterochromatin and small RNA pathways fail to transmit the response to their offspring. Also, polycomb- and H3K9me3-regulators are absolutely required in germline and in the zygote to control IGMR. Hence, I study the phenotypes of F1 embryos of males fed with high-sugar diet by analyzing genome-wide RNA-Seq. I identify changes in expression of genes normally embedded in repressive chromatin states, suggesting a role for heterochromatin in transgenerational epigenetic inheritance.

In the fourth chapter of this thesis I switch the animal model of transgenerational epigenetic inheritance to *C. elegans*. Here, in collaboration with the Lehner lab at the Centre de Regulació Genòmica (CRG, Barcelona), we show that growing worms at a high temperature for five generations increases the expression of a transgene for fourteen generations (Klosin, Casas, et al. 2017). We find that the inheritance happens in *cis* with the locus, consistent with chromatin as a potential carrier of the epigenetic information. The change in expression correlates with loss of H3K9me3 at the locus already before the onset of transcription, suggesting that the difference is inherited and not a secondary effect of higher

transcription. There are no differences observed in H3K27me3 in the locus. The repression of the transgene requires the H3K9 methyltransferase SET25, and so does the inheritance of altered expression.

Our mechanistic explanation for the model is that high temperature inhibits SET25 deposition of H3K9me3 on the transgene and this active epigenetic state is stably transmitted to the next generation. When worms are put back in normal temperature, the acquired change in expression is gradually lost after each generation until it is finally reset to normal levels consistent with a rescue of the SET25 activity. Last, we generate RNA-Seq data to study expression of endogenous genes that can behave like the transgene. Interestingly, I find that transposons that are also upregulated by temperature and repressed by SET25 also maintain memory of temperature-induced activation for several generations. Hence, transposable elements are able to keep and transmit transgenerational epigenetic modifications after environmental stimuli.

In the fifth chapter of this thesis I study heritable repression of a heterochromatic transgene in C. elegans in collaboration with the Lehner lab at the CRG. In a genome-wide screening done by members of this lab we find that RNAi targeting of many components of the core replication machinery leads to loss of silencing of the transgene. Based on this observation, we use *div-1* mutants (the gene encoding the B subunit of DNA polymerase alpha-primase) to show that impaired DNA replication during embryonic development leads to upregulation of a transgene (Klosin, Reis, et al. 2017). Since the transgene array is repressed by heterochromatin we test the effect of impaired DNA replication on animals lacking H3K27me3 and H3K39me1/2/3. Surprisingly, replication impairment upregulates expression in animals lacking any of the tested modifications, but the impact is highly reduced in animals lacking both H3K27me3 and H3K39me1/2/3. This is associated with a global reduction in these histone modifications genome-wide.

To test if the global effects of chromatin dysregulation correlate with changes in gene expression genome-wide we generate RNA-Seq data. Strikingly, I find a clear trend of endogenous genes regulated by polycomb and heterochromatin that are highly upregulated genome-wide. Since modified histones are inherited by the zygote but need to be maintained after each cell division, I quantify the loss of histones marked with H3K27me3 during embryo development and associate the lack of retention with impaired replication. The loss of repressive histone modifications causes an epigenetic state of higher expression that behaves as an epiallele and is inherited for many generations. We conclude that impaired DNA replication not only causes genetic damage but also stable epigenetic changes.

In conclusion, I show in different models that environmental stimuli can affect the epigenome of the germline and regulate gene expression in the next generation. First, I suggest that IAP insertions are enough to trigger the birth of piRNA clusters, presenting a new role of transposable elements in modifying the sperm small RNA pool. Also, I present a role of IAP insertions in regulating gene expression in embryos, suggesting that there is epigenetic inheritance of the IAP state linked to gene activation in the next generation.

Then, I also provide strong evidence of inter- and trans-generational epigenetic inheritance through the germline in flies and worms. Importantly, I have linked epigenetic inheritance to transposable elements and different epigenetic factors such as polycomb and heterochromatin. I showed that some repeat families are able to maintain acquired expression patterns for many generations and I have described inherited changes in gene expression mediated by chromatin modifications such as H3K9me3 and H3K27me3. The common epigenetic alterations in these studies are heterochromatin factors, usually bound to repetitive elements, but also relevant in gene regulation. I hypothesize that the 'junk' of the genome and its repressive mechanism play a big role in epigenetic inheritance.

Finally, although the evidence of TEI in mammals is still scarce and the mechanistic conservation in humans remains debatable, the realization that epigenetic inheritance is possible in animals and that

acute and physiological stimuli can affect the life of the progeny for their whole life has considerable implications.

To note, epigenetic alterations are common in many diseases, such as cancer, but they are also driver alterations of some important diseases in humans (reviewed in (Feinberg 2018). Hence, a perturbed inherited epigenome can have severe consequences in the life of an organism. It will therefore be very important to clarify the relationship between environmental cues, the epigenetic machinery involved, the consequent epigenetic alteration and the resulting phenotypes.

Which parts of the genome are able to maintain parental information and which are fully reset? This is one of the most interesting enigmas in the field of epigenetic inheritance.

Conclusions

Chapters 1 & 2

1. The pi-Noct piRNA cluster is the first known polymorphism in piRNA expression in mouse.

2. The pi-Noct polymorphism is associated with the presence of an IAP element in an intron of the Nocturnin gene.

3. Other IAP insertions in germline-expressed genes are also associated with piRNA production from intronic transcripts.

4. The mechanism of piRNA biogenesis from these precursors consists of a post-transcriptional processing of the Nocturnin first intron as a piRNA precursor.

5. NXF1, the nuclear RNA export factor, is required to select the Nocturnin spliced first exon for piRNA biogenesis, implicating this protein in piRNA biogenesis in mouse for the first time.

6. There is biased expression of the IAP+ allele of the Nocturnin gene in mouse embryos.

7. There is parent-of-origin specific expression of the Nocturnin gene in mouse embryos, suggesting there is maintenance of epigenetic information.

Chapter 3, 4 & 5

1. Changes in expression in the F1 after paternal high-sugar diet in flies are associated with heterochromatin-regulated genes.

2. Heterochromatin-regulated changes in gene expression after high-sugar diet are already forecast in sperm of the F0.

3. Repetitive elements maintain memory of expression after high-temperature exposure if they are repressed by heterochromatin.

4. Endogenous genes that lose heterochromatin-associated histone modifications maintain memory of upregulation.

5. Heterochromatin and transposable elements play a key role in the transmission and maintenance of differential expression in the next generation in flies and worms.

Bibliography

Allshire, Robin C., and Hiten D. Madhani. 2018. “Ten Principles of Heterochromatin Formation and Function.” *Nature Reviews. Molecular Cell Biology* 19 (4): 229–44.

Aravin, Alexei A., Ravi Sachidanandam, Deborah Bourc’his, Christopher Schaefer, Dubravka Pezic, Katalin Fejes Toth, Timothy Bestor, and Gregory J. Hannon. 2008. “A piRNA Pathway Primed by Individual Transposons Is Linked to de Novo DNA Methylation in Mice.” *Molecular Cell* 31 (6): 785–99.

Aravin, Alexei A., Ravi Sachidanandam, Angelique Girard, Katalin Fejes-Toth, and Gregory J. Hannon. 2007. “Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control.” *Science* 316 (5825): 744–47.

Aravin, Alexei, Dimos Gaidatzis, Sébastien Pfeffer, Mariana Lagos-Quintana, Pablo Landgraf, Nicola Iovino, Patricia Morris, et al. 2006. “A Novel Class of Small RNAs Bind to MILI Protein in Mouse Testes.” *Nature* 442 (7099): 203–7.

Barton, S. C., M. A. Surani, and M. L. Norris. 1984. “Role of Paternal and Maternal Genomes in Mouse Development.” *Nature* 311 (5984): 374–76.

Bell, A. C., and G. Felsenfeld. 2000. “Methylation of a CTCF-Dependent Boundary Controls Imprinted Expression of the Igf2 Gene.” *Nature* 405 (6785): 482–85.

Bogutz, Aaron B., Julie Brind’Amour, Hisato Kobayashi, Kristoffer N. Jensen, Kazuhiko Nakabayashi, Hiroo Imai, Matthew C. Lorincz, and Louis Lefebvre. n.d. “Evolution of Imprinting via Lineage-Specific Insertion of Retroviral Promoters.” https://doi.org/10.1101/723254.

Brennecke, Julius, Alexei A. Aravin, Alexander Stark, Monica Dus, Manolis Kellis, Ravi Sachidanandam, and Gregory J. Hannon. 2007. “Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila.” *Cell* 128 (6): 1089–1103.

Brennecke, Julius, Colin D. Malone, Alexei A. Aravin, Ravi Sachidanandam, Alexander Stark, and Gregory J. Hannon. 2008. “An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing.” *Science* 322

(5906): 1387–92.

Brind'Amour, Julie, Hisato Kobayashi, Julien Richard Albert, Kenjiro Shirane, Akihiko Sakashita, Asuka Kamio, Aaron Bogutz, et al. 2018. "LTR Retrotransposons Transcribed in Oocytes Drive Species-Specific and Heritable Changes in DNA Methylation." *Nature Communications* 9 (1): 3331.

Brown, S. W. 1966. "Heterochromatin." *Science* 151 (3709): 417–25.

Bull, J. J., and R. C. Vogt. 1979. "Temperature-Dependent Sex Determination in Turtles." *Science* 206 (4423): 1186–88.

Butler, Merlin G. 2009. "Genomic Imprinting Disorders in Humans: A Mini-Review." *Journal of Assisted Reproduction and Genetics* 26 (9-10): 477–86.

Capuano, Floriana, Michael Mülleder, Robert Kok, Henk J. Blom, and Markus Ralser. 2014. "Cytosine DNA Methylation Is Found in Drosophila Melanogaster but Absent in Saccharomyces Cerevisiae, Schizosaccharomyces Pombe, and Other Yeast Species." *Analytical Chemistry* 86 (8): 3697–3702.

Carmell, Michelle A., Angélique Girard, Henk J. G. van de Kant, Deborah Bourc'his, Timothy H. Bestor, Dirk G. de Rooij, and Gregory J. Hannon. 2007. "MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline." *Developmental Cell* 12 (4): 503–14.

Cavalli, Giacomo, and Edith Heard. 2019. "Advances in Epigenetics Link Genetics to the Environment and Disease." *Nature* 571 (7766): 489–99.

Cavalli, G., and R. Paro. 1998. "The Drosophila Fab-7 Chromosomal Element Conveys Epigenetic Inheritance during Mitosis and Meiosis." *Cell* 93 (4): 505–18.

———. 1999. "Epigenetic Inheritance of Active Chromatin after Removal of the Main Transactivator." *Science* 286 (5441): 955–58.

Charnier, M. 1966. "[Action of temperature on the sex ratio in the Agama agama (Agamidae, Lacertilia) embryo]." *Comptes rendus des seances de la Societe de biologie et de ses filiales* 160 (3): 620–22.

Chuong, Edward B., Nels C. Elde, and Cédric Feschotte. 2016. "Regulatory Evolution of Innate Immunity through Co-Option of Endogenous

Retroviruses." *Science* 351 (6277): 1083–87.

———. 2017. "Regulatory Activities of Transposable Elements: From Conflicts to Benefits." *Nature Reviews. Genetics* 18 (2): 71–86.

Chuong, Edward B., M. A. Karim Rumi, Michael J. Soares, and Julie C. Baker. 2013. "Endogenous Retroviruses Function as Species-Specific Enhancer Elements in the Placenta." *Nature Genetics* 45 (3): 325–29.

Concepcion, Dorothy, Kevin D. Ross, Kasey R. Hutt, Gene W. Yeo, and Bruce A. Hamilton. 2015. "Nxf1 Natural Variant E610G Is a Semi-Dominant Suppressor of IAP-Induced RNA Processing Defects." *PLoS Genetics* 11 (4): e1005123.

Crick, F. 1970. "Central Dogma of Molecular Biology." *Nature* 227 (5258): 561–63.

Czech, Benjamin, Marzia Munafò, Filippo Ciabrelli, Evelyn L. Eastwood, Martin H. Fabry, Emma Kneuss, and Gregory J. Hannon. 2018. "piRNA-Guided Genome Defense: From Biogenesis to Silencing." *Annual Review of Genetics* 52 (November): 131–57.

Darwin, Charles. 1859. "On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life /." https://doi.org/10.5962/bhl.title.68064.

Deans, Carrie, and Keith A. Maggert. 2015. "What Do You Mean, 'Epigenetic'?" *Genetics* 199 (4): 887–96.

De Fazio, Serena, Nenad Bartonicek, Monica Di Giacomo, Cei Abreu-Goodger, Aditya Sankar, Charlotta Funaya, Claude Antony, Pedro N. Moreira, Anton J. Enright, and Dónal O'Carroll. 2011. "The Endonuclease Activity of Mili Fuels piRNA Amplification That Silences LINE1 Elements." *Nature* 480 (7376): 259–63.

Deng, Wei, and Haifan Lin. 2002. "Miwi, a Murine Homolog of Piwi, Encodes a Cytoplasmic Protein Essential for Spermatogenesis." *Developmental Cell* 2 (6): 819–30.

Deniz, Özgen, Jennifer M. Frost, and Miguel R. Branco. 2019. "Regulation of Transposable Elements by DNA Modifications." *Nature Reviews. Genetics* 20 (7): 417–31.

Dennis, Cynthia, Emilie Brasset, Arpita Sarkar, and Chantal Vaury. 2016. "Export of piRNA Precursors by EJC Triggers Assembly of Cytoplasmic

Yb-Body in Drosophila." *Nature Communications* 7 (December): 13739.

Di Giacomo, Monica, Stefano Comazzetto, Harpreet Saini, Serena De Fazio, Claudia Carrieri, Marcos Morgan, Lina Vasiliauskaite, Vladimir Benes, Anton J. Enright, and Dónal O'Carroll. 2013. "Multiple Epigenetic Mechanisms and the piRNA Pathway Enforce LINE1 Silencing during Adult Spermatogenesis." *Molecular Cell* 50 (4): 601–8.

Eckersley-Maslin, Melanie A., Celia Alda-Catalinas, and Wolf Reik. 2018. "Dynamics of the Epigenetic Landscape during the Maternal-to-Zygotic Transition." *Nature Reviews. Molecular Cell Biology* 19 (7): 436–50.

Elgin, Sarah C. R., and Gunter Reuter. 2013. "Position-Effect Variegation, Heterochromatin Formation, and Gene Silencing in Drosophila." *Cold Spring Harbor Perspectives in Biology* 5 (8): a017780.

Ellison, Christopher E., and Doris Bachtrog. 2013. "Dosage Compensation via Transposable Element Mediated Rewiring of a Regulatory Network." *Science* 342 (6160): 846–50.

Ernst, Christina, Duncan T. Odom, and Claudia Kutter. 2017. "The Emergence of piRNAs against Transposon Invasion to Preserve Mammalian Genome Integrity." *Nature Communications* 8 (1): 1411.

Faulkner, Geoffrey J., Yasumasa Kimura, Carsten O. Daub, Shivangi Wani, Charles Plessy, Katharine M. Irvine, Kate Schroder, et al. 2009. "The Regulated Retrotransposon Transcriptome of Mammalian Cells." *Nature Genetics* 41 (5): 563–71.

Feinberg, Andrew P. 2018. "The Key Role of Epigenetics in Human Disease Prevention and Mitigation." *The New England Journal of Medicine* 378 (14): 1323–34.

Ferguson, M. W., and T. Joanen. 1982. "Temperature of Egg Incubation Determines Sex in Alligator Mississippiensis." *Nature* 296 (5860): 850–53.

Fort, Alexandre, Kosuke Hashimoto, Daisuke Yamada, Md Salimullah, Chaman A. Keya, Alka Saxena, Alessandro Bonetti, et al. 2014. "Deep Transcriptome Profiling of Mammalian Stem Cells Supports a Regulatory Role for Retrotransposons in Pluripotency Maintenance." *Nature Genetics* 46 (6): 558–66.

Gaydos, Laura J., Wenchao Wang, and Susan Strome. 2014. "Gene

Repression. H3K27me and PRC2 Transmit a Memory of Repression across Generations and during Development." *Science* 345 (6203): 1515–18.

Ge, Chutian, Jian Ye, Ceri Weber, Wei Sun, Haiyan Zhang, Yingjie Zhou, Cheng Cai, Guoying Qian, and Blanche Capel. 2018. "The Histone Demethylase KDM6B Regulates Temperature-Dependent Sex Determination in a Turtle Species." *Science* 360 (6389): 645–48.

Girard, Angélique, Ravi Sachidanandam, Gregory J. Hannon, and Michelle A. Carmell. 2006. "A Germline-Specific Class of Small RNAs Binds Mammalian Piwi Proteins." *Nature* 442 (7099): 199–202.

Grentzinger, Thomas, Claudia Armenise, Christine Brun, Bruno Mugat, Vincent Serrano, Alain Pelisson, and Séverine Chambeyron. 2012. "piRNA-Mediated Transgenerational Inheritance of an Acquired Trait." *Genome Research* 22 (10): 1877–88.

Grivna, Shane T., Ergin Beyret, Zhong Wang, and Haifan Lin. 2006. "A Novel Class of Small RNAs in Mouse Spermatogenic Cells." *Genes & Development* 20 (13): 1709–14.

Gunawardane, Lalith S., Kuniaki Saito, Kazumichi M. Nishida, Keita Miyoshi, Yoshinori Kawamura, Tomoko Nagami, Haruhiko Siomi, and Mikiko C. Siomi. 2007. "A Slicer-Mediated Mechanism for Repeat-Associated siRNA 5' End Formation in Drosophila." *Science* 315 (5818): 1587–90.

Han, Bo W., Wei Wang, Chengjian Li, Zhiping Weng, and Phillip D. Zamore. 2015. "Noncoding RNA. piRNA-Guided Transposon Cleavage Initiates Zucchini-Dependent, Phased piRNA Production." *Science* 348 (6236): 817–21.

Heard, Edith, and Robert A. Martienssen. 2014. "Transgenerational Epigenetic Inheritance: Myths and Mechanisms." *Cell* 157 (1): 95–109.

Heijden, G. W. van der, A. A. H. A. Derijck, L. Ramos, M. Giele, J. van der Vlag, and P. de Boer. 2006. "Transmission of Modified Nucleosomes from the Mouse Male Germline to the Zygote and Subsequent Remodeling of Paternal Chromatin." *Developmental Biology* 298 (2): 458–69.

Heintzman, Nathaniel D., Gary C. Hon, R. David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F. Harp, Zhen Ye, et al. 2009. "Histone Modifications at Human Enhancers Reflect Global Cell-Type-Specific Gene Expression." *Nature* 459 (7243): 108–12.

Hon, Gary C., R. David Hawkins, and Bing Ren. 2009. “Predictive Chromatin Signatures in the Mammalian Genome.” *Human Molecular Genetics* 18 (R2): R195–201.

Inoue, Azusa, Lan Jiang, Falong Lu, Tsukasa Suzuki, and Yi Zhang. 2017. “Maternal H3K27me3 Controls DNA Methylation-Independent Imprinting.” *Nature* 547 (7664): 419–24.

Kazachenka, Anastasiya, Tessa M. Bertozzi, Marcela K. Sjoberg-Herrera, Nic Walker, Joseph Gardner, Richard Gunning, Elena Pahita, Sarah Adams, David Adams, and Anne C. Ferguson-Smith. 2018. “Identification, Characterization, and Heritability of Murine Metastable Epialleles: Implications for Non-Genetic Inheritance.” *Cell* 175 (6): 1717.

Klosin, Adam, Eduard Casas, Cristina Hidalgo-Carcedo, Tanya Vavouri, and Ben Lehner. 2017. “Transgenerational Transmission of Environmental Information in.” *Science* 356 (6335): 320–23.

Klosin, Adam, Kadri Reis, Cristina Hidalgo-Carcedo, Eduard Casas, Tanya Vavouri, and Ben Lehner. 2017. “Impaired DNA Replication Derepresses Chromatin and Generates a Transgenerationally Inherited Epigenetic Memory.” *Science Advances* 3 (8): e1701143.

Kojima-Kita, Kanako, Satomi Kuramochi-Miyagawa, Ippei Nagamori, Narumi Ogonuki, Atsuo Ogura, Hidetoshi Hasuwa, Takashi Akazawa, Norimitsu Inoue, and Toru Nakano. 2016. “MIWI2 as an Effector of DNA Methylation and Gene Silencing in Embryonic Male Germ Cells.” *Cell Reports* 16 (11): 2819–28.

Kolasinska-Zwierz, Paulina, Thomas Down, Isabel Latorre, Tao Liu, X. Shirley Liu, and Julie Ahringer. 2009. “Differential Chromatin Marking of Introns and Expressed Exons by H3K36me3.” *Nature Genetics* 41 (3): 376–81.

Kruse, Kai, Noelia Diaz, Rocio Enriquez-Gasca, Xavier Gaume, Maria-Elena Torres-Padilla, and Juan M. Vaquerizas. n.d. “Transposable Elements Drive Reorganisation of 3D Chromatin during Early Embryogenesis.” https://doi.org/10.1101/523712.

Kunarso, Galih, Na-Yu Chia, Justin Jeyakani, Catalina Hwang, Xinyi Lu, Yun-Shen Chan, Huck-Hui Ng, and Guillaume Bourque. 2010. “Transposable Elements Have Rewired the Core Regulatory Network of Human Embryonic Stem Cells.” *Nature Genetics* 42 (7): 631–34.

Kuramochi-Miyagawa, Satomi, Tohru Kimura, Takashi W. Ijiri, Taku Isobe, Noriko Asada, Yukiko Fujita, Masahito Ikawa, et al. 2004. "Mili, a Mammalian Member of Piwi Family Gene, Is Essential for Spermatogenesis." *Development* 131 (4): 839–49.

Kuramochi-Miyagawa, Satomi, Toshiaki Watanabe, Kengo Gotoh, Yasushi Totoki, Atsushi Toyoda, Masahito Ikawa, Noriko Asada, et al. 2008. "DNA Methylation of Retrotransposon Genes Is Regulated by Piwi Family Members MILI and MIWI2 in Murine Fetal Testes." *Genes & Development* 22 (7): 908–17.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, et al. 2001. "Initial Sequencing and Analysis of the Human Genome." *Nature* 409 (6822): 860–921.

Lane, Natasha, Wendy Dean, Sylvia Erhardt, Petra Hajkova, Azim Surani, Jörn Walter, and Wolf Reik. 2003. "Resistance of IAPs to Methylation Reprogramming May Provide a Mechanism for Epigenetic Inheritance in the Mouse." *Genesis* 35 (2): 88–93.

Lau, N. C., A. G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakano, D. P. Bartel, and R. E. Kingston. 2006. "Characterization of the piRNA Complex from Rat Testes." *Science* 313 (5785): 363–67.

Le Thomas, Adrien, Evelyn Stuwe, Sisi Li, Jiamu Du, Georgi Marinov, Nikolay Rozhkov, Yung-Chia Ariel Chen, et al. 2014. "Transgenerationally Inherited piRNAs Trigger piRNA Biogenesis by Changing the Chromatin of piRNA Clusters and Inducing Precursor Processing." *Genes & Development* 28 (15): 1667–80.

Li, Jingfeng, Keiko Akagi, Yongjun Hu, Anna L. Trivett, Christopher J. W. Hlynialuk, Deborah A. Swing, Natalia Volfovsky, et al. 2012. "Mouse Endogenous Retroviruses Can Trigger Premature Transcriptional Termination at a Distance." *Genome Research* 22 (5): 870–84.

Luedi, Philippe P., Alexander J. Hartemink, and Randy L. Jirtle. 2005. "Genome-Wide Prediction of Imprinted Murine Genes." *Genome Research* 15 (6): 875–84.

Luger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. "Crystal Structure of the Nucleosome Core Particle at 2.8 A Resolution." *Nature* 389 (6648): 251–60.

Macfarlan, Todd S., Wesley D. Gifford, Shawn Driscoll, Karen Lettieri, Helen

M. Rowe, Dario Bonanomi, Amy Firth, Oded Singer, Didier Trono, and Samuel L. Pfaff. 2012. “Embryonic Stem Cell Potency Fluctuates with Endogenous Retrovirus Activity.” *Nature* 487 (7405): 57–63.

Mcclintock, B. 1956. “Controlling Elements and the Gene.” *Cold Spring Harbor Symposia on Quantitative Biology* 21: 197–216.

McClintock, Barbara. 1961. “Some Parallels Between Gene Control Systems in Maize and in Bacteria.” *The American Naturalist*. https://doi.org/10.1086/282188.

McClintock, Barbara, and Others. 1958. “The Suppressor-Mutator System of Control of Gene Action in Maize.” *Year Book - Carnegie Institution of Washington* 57: 415–29.

McGrath, J., and D. Solter. 1984. “Completion of Mouse Embryogenesis Requires Both the Maternal and Paternal Genomes.” *Cell* 37 (1): 179–83.

Mendel, Gregor. 1865. “Versuche über Plflanzen-Hybriden.” *Verhandlungen Des Naturforschenden Ver-Eines in Brünn* IV: 3–47.

Mohn, Fabio, Dominik Handler, and Julius Brennecke. 2015. “Noncoding RNA. piRNA-Guided Slicing Specifies Transcripts for Zucchini-Dependent, Phased piRNA Biogenesis.” *Science* 348 (6236): 812–17.

Moore, Rebecca S., Rachel Kaletsky, and Coleen T. Murphy. 2019. “Piwi/PRG-1 Argonaute and TGF-β Mediate Transgenerational Learned Pathogenic Avoidance.” *Cell* 177 (7): 1827–41.e12.

Morgan, H. D., H. G. Sutherland, D. I. Martin, and E. Whitelaw. 1999. “Epigenetic Inheritance at the Agouti Locus in the Mouse.” *Nature Genetics* 23 (3): 314–18.

Mouse Genome Sequencing Consortium. 2002. “Initial Sequencing and Comparative Analysis of the Mouse Genome.” *Nature* 420 (6915): 520–62.

Muller, Hans J. 1930. “Types of Visible Variations Induced by X-Rays inDrosophila.” *Journal of Genetics* 22 (3): 299–334.

Nirenberg, M. W., and J. H. Matthaei. 1961. “The Dependence of Cell-Free Protein Synthesis in E. Coli upon Naturally Occurring or Synthetic Polyribonucleotides.” *Proceedings of the National Academy of Sciences of the United States of America* 47 (October): 1588–1602.

Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. “DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for de Novo Methylation and Mammalian Development.” *Cell* 99 (3): 247–57.

Ospina-Alvarez, Natalia, and Francesc Piferrer. 2008. “Temperature-Dependent Sex Determination in Fish Revisited: Prevalence, a Single Sex Ratio Response Pattern, and Possible Effects of Climate Change.” *PloS One* 3 (7): e2837.

Öst, Anita, Adelheid Lempradl, Eduard Casas, Melanie Weigert, Theodor Tiko, Merdin Deniz, Lorena Pantano, et al. 2014. “Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity.” *Cell* 159 (6): 1352–64.

Ozata, Deniz M., Ildar Gainetdinov, Ansgar Zoch, Dónal O’Carroll, and Phillip D. Zamore. 2019. “PIWI-Interacting RNAs: Small RNAs with Big Functions.” *Nature Reviews. Genetics* 20 (2): 89–108.

Perez, Marcos Francisco, and Ben Lehner. 2019. “Intergenerational and Transgenerational Epigenetic Inheritance in Animals.” *Nature Cell Biology* 21 (2): 143–51.

Posner, Rachel, Itai Antoine Toker, Olga Antonova, Ekaterina Star, Sarit Anava, Eran Azmon, Michael Hendricks, Shahar Bracha, Hila Gingold, and Oded Rechavi. 2019. “Neuronal Small RNAs Control Behavior Transgenerationally.” *Cell* 177 (7): 1814–26.e15.

Rakyan, Vardhman K., Marnie E. Blewitt, Riki Druker, Jost I. Preis, and Emma Whitelaw. 2002. “Metastable Epialleles in Mammals.” *Trends in Genetics*. https://doi.org/10.1016/s0168-9525(02)02709-9.

Reik, W., W. Dean, and J. Walter. 2001. “Epigenetic Reprogramming in Mammalian Development.” *Science* 293 (5532): 1089–93.

Reuter, Michael, Philipp Berninger, Shinichiro Chuma, Hardik Shah, Mihoko Hosokawa, Charlotta Funaya, Claude Antony, Ravi Sachidanandam, and Ramesh S. Pillai. 2011. “Miwi Catalysis Is Required for piRNA Amplification-Independent LINE1 Transposon Silencing.” *Nature* 480 (7376): 264–67.

Rowe, Helen M., Johan Jakobsson, Daniel Mesnard, Jacques Rougemont, Séverine Reynard, Tugce Aktas, Pierre V. Maillard, et al. 2010. “KAP1 Controls Endogenous Retroviruses in Embryonic Stem Cells.” *Nature* 463 (7278): 237–40.

Sasaki, Hiroyuki, and Yasuhisa Matsui. 2008. “Epigenetic Events in Mammalian Germ-Cell Development: Reprogramming and beyond.” *Nature Reviews. Genetics* 9 (2): 129–40.

Schultz, J. 1936. “Variegation in Drosophila and the Inert Chromosome Regions.” *Proceedings of the National Academy of Sciences of the United States of America* 22 (1): 27–33.

Sharma, Upasna, Fengyun Sun, Colin C. Conine, Brian Reichholf, Shweta Kukreja, Veronika A. Herzog, Stefan L. Ameres, and Oliver J. Rando. 2018. “Small RNAs Are Trafficked from the Epididymis to Developing Mammalian Sperm.” *Developmental Cell* 46 (4): 481–94.e6.

Simpson, V. J., T. E. Johnson, and R. F. Hammen. 1986. “Caenorhabditis Elegans DNA Does Not Contain 5-Methylcytosine at Any Time during Development or Aging.” *Nucleic Acids Research* 14 (16): 6711–19.

Slotkin, R. Keith, and Robert Martienssen. 2007. “Transposable Elements and the Epigenetic Regulation of the Genome.” *Nature Reviews. Genetics* 8 (4): 272–85.

Sotero-Caio, Cibele G., Roy N. Platt 2nd, Alexander Suh, and David A. Ray. 2017. “Evolution and Diversity of Transposable Elements in Vertebrate Genomes.” *Genome Biology and Evolution* 9 (1): 161–77.

Srivastava, Deepak, and Natalie DeWitt. 2016. “In Vivo Cellular Reprogramming: The Next Generation.” *Cell* 166 (6): 1386–96.

Surani, M. A., S. C. Barton, and M. L. Norris. 1984. “Development of Reconstituted Mouse Eggs Suggests Imprinting of the Genome during Gametogenesis.” *Nature* 308 (5959): 548–50.

Tucci, Valter, Anthony R. Isles, Gavin Kelsey, Anne C. Ferguson-Smith, and Erice Imprinting Group. 2019. “Genomic Imprinting and Physiological Processes in Mammals.” *Cell* 176 (5): 952–65.

Vourekas, Anastassios, Qi Zheng, Panagiotis Alexiou, Manolis Maragkakis, Yohei Kirino, Brian D. Gregory, and Zissimos Mourelatos. 2012. “Mili and Miwi Target RNA Repertoire Reveals piRNA Biogenesis and Function of Miwi in Spermiogenesis.” *Nature Structural & Molecular Biology* 19 (8): 773–81.

Waddington, C. H. 1942. “CANALIZATION OF DEVELOPMENT AND THE INHERITANCE OF ACQUIRED CHARACTERS.” *Nature*.

https://doi.org/10.1038/150563a0.

———. 1953. “Genetic Assimilation of an Acquired Character.” *Evolution*. https://doi.org/10.2307/2405747.

Walsh, C. P., J. R. Chaillet, and T. H. Bestor. 1998. “Transcription of IAP Endogenous Retroviruses Is Constrained by Cytosine Methylation.” *Nature Genetics* 20 (2): 116–17.

Wang, J., J. Mager, Y. Chen, E. Schneider, J. C. Cross, A. Nagy, and T. Magnuson. 2001. “Imprinted X Inactivation Maintained by a Mouse Polycomb Group Gene.” *Nature Genetics* 28 (4): 371–75.

Wang, Zhibin, Chongzhi Zang, Jeffrey A. Rosenfeld, Dustin E. Schones, Artem Barski, Suresh Cuddapah, Kairong Cui, et al. 2008. “Combinatorial Patterns of Histone Acetylations and Methylations in the Human Genome.” *Nature Genetics* 40 (7): 897–903.

Watanabe, Toshiaki, Shin-Ichi Tomizawa, Kohzoh Mitsuya, Yasushi Totoki, Yasuhiro Yamamoto, Satomi Kuramochi-Miyagawa, Naoko Iida, et al. 2011. “Role for piRNAs and Noncoding RNA in de Novo DNA Methylation of the Imprinted Mouse Rasgrf1 Locus.” *Science* 332 (6031): 848–52.

Watson, J. D., and F. H. Crick. 1953. “Molecular Structure of Nucleic Acids; a Structure for Deoxyribose Nucleic Acid.” *Nature* 171 (4356): 737–38.

Weismann, August. 1893. “The Germ-Plasm: A Theory of Heredity. Translated by W. Newton Parker and Harriet Rönnfeldt.” https://doi.org/10.5962/bhl.title.25196.

Wu, Pei-Hsuan, Yu Fu, Katharine Cecchini, Deniz M. Özata, Zhiping Weng, and Phillip D. Zamore. 2018. “An Evolutionarily Conserved piRNA-Producing Locus Required for Male Mouse Fertility.” *bioRxiv*. https://doi.org/10.1101/386201.

Yoder, J. A., C. P. Walsh, and T. H. Bestor. 1997. “Cytosine Methylation and the Ecology of Intragenomic Parasites.” *Trends in Genetics: TIG* 13 (8): 335–40.

Zhang, Yanxiao, Ting Li, Sebastian Preissl, Maria Luisa Amaral, Jonathan D. Grinstein, Elie N. Farah, Eugin Destici, et al. 2019. “Transcriptionally Active HERV-H Retrotransposons Demarcate Topologically Associating Domains in Human Pluripotent Stem Cells.” *Nature Genetics* 51 (9):

1380–88.

Zink, D., and R. Paro. 1995. “Drosophila Polycomb-Group Regulated Chromatin Inhibits the Accessibility of a Trans-Activator to Its Target DNA.” *The EMBO Journal* 14 (22): 5660–71.

Annex

Chapter 3: Paternal diet defines offspring chromatin state and intergenerational obesity

Chapter 4: Transgenerational transmission of environmental information in C. elegans

Chapter 5: Impaired DNA replication derepresses chromatin and generates a transgenerationally inherited epigenetic memory

Cell

Article

Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity

Anita Öst,[1,6,8,*] Adelheid Lempradl,[1,8] Eduard Casas,[2,3] Melanie Weigert,[1] Theodor Tiko,[1] Merdin Deniz,[1] Lorena Pantano,[2] Ulrike Boenisch,[1] Pavel M. Itskov,[7] Marlon Stoeckius,[4] Marius Ruf,[1] Nikolaus Rajewsky,[4] Gunter Reuter,[5] Nicola Iovino,[1] Carlos Ribeiro,[7] Mattias Alenius,[6] Steffen Heyne,[1] Tanya Vavouri,[2,3] and J. Andrew Pospisilik[1,*]

[1]Max Planck Institute of Immunobiology and Epigenetics, Stuebeweg 51, 79108 Freiburg, Germany
[2]Institut de Medicina Predictiva i Personalitzada del Càncer, Can Ruti Campus, Ctra. de Can Ruti, Camí de les Escoles s/n, 08916 Badalona, Barcelona, Spain
[3]Josep Carreras Leukaemia Research Institute (IJC), ICO-Hospital GermansTrias i Pujol, 08916 Badalona, Barcelona, Spain
[4]Systems Biology of Gene Regulatory Elements, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany
[5]Institute of Biology, Developmental Genetics, Martin Luther University Halle, 06120 Halle, Germany
[6]Department of Clinical and Experimental Medicine, Linkoping University, 58183 Linkoping, Sweden
[7]Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal
[8]Co-first author
*Correspondence: anita.ost@liu.se (A.Ö.), pospisilik@ie-freiburg.mpg.de (J.A.P.)
http://dx.doi.org/10.1016/j.cell.2014.11.005

SUMMARY

The global rise in obesity has revitalized a search for genetic and epigenetic factors underlying the disease. We present a *Drosophila* model of paternal-diet-induced *i*ntergenerational *m*etabolic *r*eprogramming (IGMR) and identify genes required for its encoding in offspring. Intriguingly, we find that as little as 2 days of dietary intervention in fathers elicits obesity in offspring. Paternal sugar acts as a physiological suppressor of variegation, desilencing chromatin-state-defined domains in both mature sperm and in offspring embryos. We identify requirements for H3K9/K27me3-dependent reprogramming of metabolic genes in two distinct germline and zygotic windows. Critically, we find evidence that a similar system may regulate obesity susceptibility and phenotype variation in mice and humans. The findings provide insight into the mechanisms underlying intergenerational metabolic reprogramming and carry profound implications for our understanding of phenotypic variation and evolution.

INTRODUCTION

Global incidence of obesity is approaching 1 billion humans. Though poorly understood, parental and fetal nutritional states have been shown to generate reproducible offspring phenotypes, including obesity. Studies in multiple model organisms have been used to examine intergenerational metabolic effects (Braunschweig et al., 2012; Ozanne et al., 1999; Morgan et al., 1999; Buescher et al., 2013; Rechavi et al., 2014). Maternal and paternal induction of intergenerational responses have been reported, and a variety of macronutrient and timing interventions have been used, including short- and long-term fasting (Anderson et al., 2006), calorie restriction (Blondeau et al., 2002; Jimenez-Chillaron et al., 2009), and modulation of dietary protein (Ozanne et al., 1999), fat (Gniuli et al., 2008; Dunn and Bale, 2009), and methyl-donor content (Wolff et al., 1998; Waterland et al., 2006; reviewed in Daxinger and Whitelaw, 2012; Patti, 2013). Of note, although not understood, divergent physiological extremes can prompt similar offspring phenotypes, so called "U-shaped" responses.

Intergenerational effects transmitted via the male germline have received recent attention. Because father-to-offspring transmission excludes difficult to control oocyte and gestational effects, mechanistic dissections are simplified. Studies have demonstrated paternal transmission of tumor susceptibility (Anway et al., 2005; Xing et al., 2007), of heat-shock-induced epigenetic memory (Seong et al., 2011), of olfaction-dependent behavioral and neural phenotypes (Dias and Ressler, 2014), and of metabolic control (Anderson et al., 2006; Fullston et al., 2013; Carone et al., 2010; Ng et al., 2010; reviewed in Rando, 2012).

Mechanistically, imprinting, altered DNA methylation, histone modifications, and noncoding RNA transcripts have been implicated in inter/transgenerational phenotype transmission. Adiposity of genetically identical agouti mouse siblings correlates with IAP DNA methylation (Morgan et al., 1999); DNA methylation correlates with endocrine disruptor and nutrient induced inter/transgenerational phenotypes (Anway et al., 2005; Carone et al., 2010; Radford et al., 2014), and there is evidence of RNA-dependent transmission (Gapp et al., 2014; Rechavi et al., 2014; Kiani et al., 2013; Rassoulzadegan et al., 2006). In *C. elegans* and *Drosophila*, research has focused on small noncoding RNAs and chromatin organization (Seong et al., 2011; Shirayama et al., 2012; Lee et al., 2012; Greer et al., 2011; Ashe et al., 2012). Despite these advances, however, our understanding of the initiation, transmission, and stabilization of trans/intergenerational phenotypes remains largely a black box.

Here, we present a *Drosophila* model of paternal *i*ntergenerational *m*etabolic *r*eprogramming (IGMR) and identify

germline and zygotic gene networks that are necessary for its manifestation. Mechanistically, paternal sugar modifies offspring chromatin state and transcription in a *Polycomb-*, *E(z)-*, *SetDB1-*, *Su(var)3-9-*, and *HP1*-sensitive manner. Intriguingly, these changes are forecast in the sperm. Data from highly defined human and mouse obesity cohorts suggest that these processes are conserved. These data provide evidence for a conserved chromatin-state-encoded program that defines phenotypic variation and thus carry profound implications for our understanding of phenotypic diversity and evolution.

RESULTS

A *Drosophila* Model of Intergenerational Metabolic State Control

We sought to understand whether normal fluctuations in diet might impact next-generation phenotypes. We chose to focus on the male germline and, for simplicity, on male progeny.

To minimize genetic variation, we performed ten generations of single-fly, brother-sister inbreeding of our population inbred w^{1118} *Drosophila melanogaster* strain. To identify an optimal dietary intervention for P0 fathers, we challenged 4- to 5-day-old males with progressively increasing dietary sugar and protein and assessed whole-fly fat storage after 2 days. Whereas dietary protein showed minimal effects (Figure 1A; horizontal axis), dietary sugar evoked a 3-fold increase in whole-fly triglyceride storage (Figure 1A; vertical axis). Of note, the sugar concentrations used approximate natural food sources (ripened banana ~300 g/l). These responses agreed with published data (Skorupa et al., 2008) and highlighted the rapid metabolic regulatory potential of dietary sugar in the fly.

To test for intergenerational effects, we repeated the experiment, this time varying only sugar, and mated the males to standardized w^{1118} female virgins (Figure 1B). After 2 hr of mating, females were left to lay eggs for 10 hr, removed, and the F1 offspring were left to develop unimpeded. Importantly, ancestral (more than ten generations), parental, and F1 generations were highly controlled with respect to male:female mating ratio, larval density, diet, and environmental conditions. One week after eclosure, adult male offspring were weighed and sacrificed, and triglyceride levels were measured in whole-fly lysates. Interestingly, although the paternal intervention showed no effect on F1 kept on normal food (Figure 1C, top, open circles), adult progeny fed an obesogenic high-sugar diet exhibited a U-shaped obesity response (Figure 1C, top closed circles) with low- and high-sugar sired individuals showing exaggerated triglycerides (Figure 1C and 1D). This phenotype was significant by both ANOVA and comparison of linear versus polynomial regressions. Thus, paternal sugar outside of the physiological optimum alters metabolic control in the F1.

The IGMR phenotype comprised two features. F1 offspring body weight increased with paternal sugar (Figure 1C, middle), and weight-normalized triglyceride levels increased toward both paternal extremes (Figure 1C and 1D). Metabolic phenotyping revealed that obesity-susceptible IGMR progeny exhibited increased adipose area (Figure 1E and 1F) and lipid droplet size (Figure 1E and Figure S1A available online). Measures of feeding behavior showed a tendency toward increased food intake (Figure 1G). Together with increased starvation sensitivity (Figure S1D), unaltered activity and CO_2 production measures (Figures S1B and S1C), as well as unaltered trehalose and glucose levels (Figures S1E and S1F), these findings suggest that the observed excess triglyceride reserves resulted from poor lipid store mobilization and possibly hyperphagia. Importantly, we found no evidence of altered ecclosure timing (Figure 1H) or wing size (Figure 1I, top), or in F1 offspring number (Figure 1I, middle) or male:female ratio per brood (Figure 1I, bottom). Thus, acute paternal dietary sugar reprograms offspring metabolism, leaving growth and development intact. Our data conclusively provide evidence that acute paternal diet reprograms offspring metabolism in *Drosophila*.

Paternal IGMR Is Rapid, Stable, and Stress Sensitive

The short 2 day intervention implied that mature *Drosophila* sperm are capable of continuously transmitting environmental cues to their offspring. To understand the minimum dietary intervention required to elicit paternal IGMR, w^{1118} males were subject to dietary interventions lasting 1, 2, 5, or 7 days prior to mating (Figure 2A). Measurements of adiposity indicated a maximal F1 phenotypic response within just 2 days of paternal challenge (Figures 2B and 2C), suggesting that paternal IGMR might be detectable within a single day. Female flies store sperm upon mating, allowing them to fertilize eggs for days or even weeks after a single insemination event. We asked how stable the IGMR phenotype would be with sperm storage. Males were subject to a 2 day paternal dietary intervention and mated, and the recipient females were allowed to lay three consecutive batches of fertilized eggs over 60 hr. Importantly, offspring of all three consecutive batches exhibited clear U-shaped obesity phenotypes (Figures 2D and 2E), indicating that IGMR is stable with sperm storage. Further, heat shock is known to modulate epigenetically controlled phenotypes, even across generations (Seong et al., 2011). Interestingly, a 1 hr heat shock of diet-treated fathers immediately before mating completely abrogated the paternal IGMR obesity response (Figures 2F and 2G). This is consistent with the requirement for highly controlled environmental conditions (sound, odor, vibration) when using this model. Finally, we found no evidence that the IGMR phenotype is transmitted to subsequent generations (Figure 2H). Thus, paternal IGMR is rapid, stable, and acutely stress sensitive.

Paternal Sugar Alters Offspring Heterochromatin

Paternal IGMR appeared phenotypically "silent" through the complexities of development. We therefore hypothesized that the phenotype was encoded in chromatin. Position-effect variegation (PEV) is a genetic phenomenon that has been used as a quantitative readout of locus-specific chromatin state silencing in vivo. The most common PEV reporters in *Drosophila* reflect chromatin desilencing through increased expression of a red-eye-pigment-coding reporter gene. Screening a library of PEV strains, Phalke and colleagues recently defined at least five functionally distinct chromatin silencing subtypes in the living fly (Phalke et al., 2009). Using identical or comparable PEV lines, we tested whether paternal dietary sugar could alter offspring eye color and thus stably alter chromatin state in F1 (Figures 3A–3E). We observed no overt effect of paternal diet on offspring

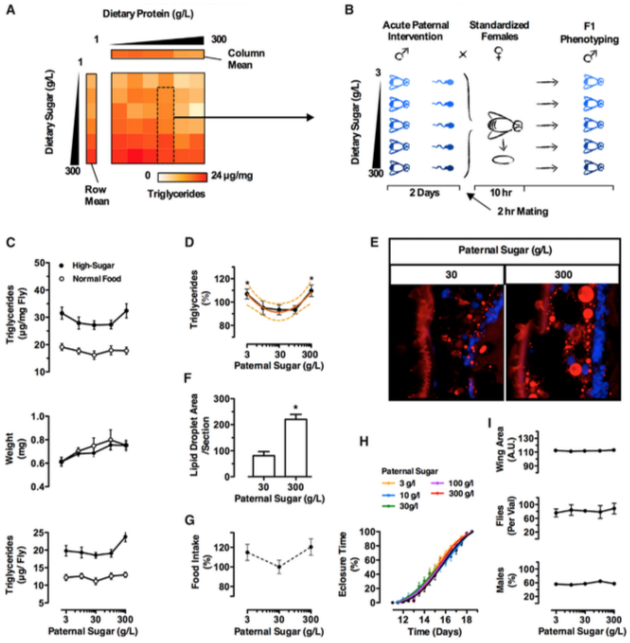


Figure 1. A Fly Model of Paternally Induced Obesity

(A) Triglyceride levels in founder males after 2 days of diet intervention with increasing sugar (sucrose) and protein (soy).

(B) Schematic of the IGMR experimental design.

(C) Triglycerides (top), body weight (middle), and weight normalized triglycerides (bottom) of F1 males raised on normal (open circles) and on high-sugar food (closed circles).

(D) Binomial regression of F1 male weight normalized triglycerides (95% CI, $p < 0.01$).

(E) Representative section of anterior fly fat body stained with oil red O and DAPI.

(F) Lipid droplet area/section.

(G) Food intake of F1 males by CAFE assay.

(H) Ecclosure timing of F1 offspring as percent of total.

(I) (Top) Relative wing area of F1 males. (Middle) Total number of ecclosed offspring per vial. (Bottom) Number of male offspring per vial (% of total flies).

Results are mean ± SEM (*$p < 0.05$) of n = 3–8 experiments each with multiple replicates. See also Figure S1.

PEV in four of the lines tested (A_{480}; Figures 3A–3D), including reporters for telomeric (ChrX; *HA-1902*) (Figure 3A), retrotransposon-type (Chr3R; *HA-1992*) (Figure 3B), pericentric (Chr4;*39c-12*) (Figure 3C), and repeat-associated chromatin (Chr2:3;*92E*) (Figure 3D). Notably, all four lines generated U-shaped paternal IGMR obesity (Figure 3F). Thus, IGMR occurs on independent genetic backgrounds and leaves HA-1902-, HA-1992-, 39c-12-, and 92E-type chromatin largely unaltered.

Intriguingly, when testing w^{m4h}, a reporter for peri-centric heterochromatin on ChrX, we observed a reproducible U-shaped intergenerational eye color phenotype (Figures 3E and 3G). In support of a mechanistic link between the IGMR obesity and

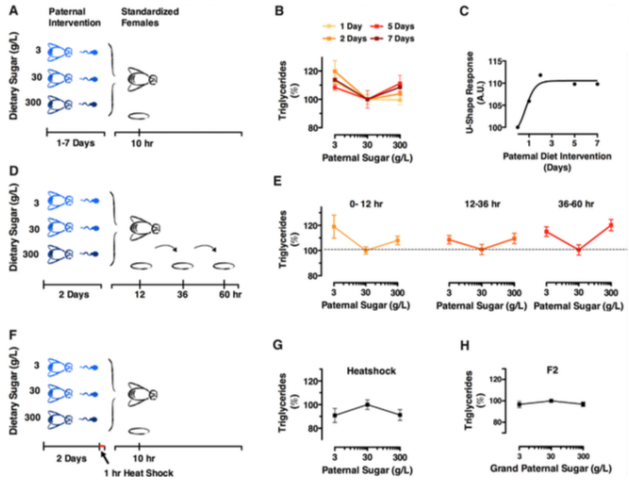


Figure 2. Acute Paternal Nutritional State Is Reflected in Offspring Obesity

(A and B) (A) Schematic and (B) offspring adiposity from tests of progressively increasing paternal dietary intervention.

(C) Mean change in offspring adiposity (Δtriglycerides/weight) for low- and high-sugar-sired adult males relative to medium sugar for each time point. Least square curve fitting (slope = −1, R^2 = 0.95).

(D and (E) (D) Schematic and (E) F1 male adult adiposity from tests of consecutive offspring cohorts from the same mating event. After mating, females were kept on standard food, and three consecutive batches of embryos were collected and assessed at adulthood.

(F) and (G) (F) Schematic and (G) offspring adiposity for tests of stress sensitivity. A 1 hr 37 degree heat-shock was applied to males just before mating.

(H) F2 adult male adiposity. F1 males were kept on standard food prior to mating.

Results are mean ± SEM of n = 3–8 experiments each with multiple replicates.

PEV results, triglyceride accumulation and eye color correlated positively; redder-eyed flies were more obese (Figure 3H). No correlation was observed in the remaining four strains (data not shown). These data show that acute paternal diet targets select chromatin subtypes in offspring.

High Paternal Sugar Controls Heterochromatin-Embedded Gene Expression

At this point, we focused on medium- versus high-sugar IGMR and tested whether IGMR affected all or only select individuals in the population. Measuring pigment from single w^{m4h} fly heads as a direct readout of the IGMR response, we observed that paternal IGMR red-shifted the entire distribution (Figures 4A and 4B). Thus, high paternal sugar induces w^{m4h} desilencing population-wide, indicating that each paternal gamete carries an equivalent intergenerational signal.

Next, we performed rRNA-depleted RNA-sequencing of hand-picked stage 17 embryo F1 offspring from medium- and high-sugar challenged fathers (Pearson corr. = 0.97, ~15 million reads/sample; Figure 4C). In support of a selective chromatin state desilencing mechanism, gene expression broadly increased, with many more up- than downregulated transcripts. Sixty-eight protein-coding genes were significantly upregulated in high-sugar sired embryos (mean ΔFPKM = 54.9) and only ten downregulated (mean ΔFPKM = −7.0; Figure 4D and Table S1). Of note, upregulated transcripts tended to be genes highly expressed during late embryo and early larval stages, including 27 (40%) related to biogenesis of the sugar-based cuticle. Of the remaining 42 genes, 30 were of unknown function, 5 had peptidase activity, and interestingly, 4 were metabolic genes, including fatty acyl-CoA reductase and fatty acid elongase.

Analysis using gene set enrichment analysis (GSEA) revealed two clearly upregulated clusters containing *chitin and cuticle constituent* and *mitochondrial and primary energy metabolism* pathways (Figures 4E and S2A and Table S2). Included and consistent with the heightened adiposity of IGMR, pathways

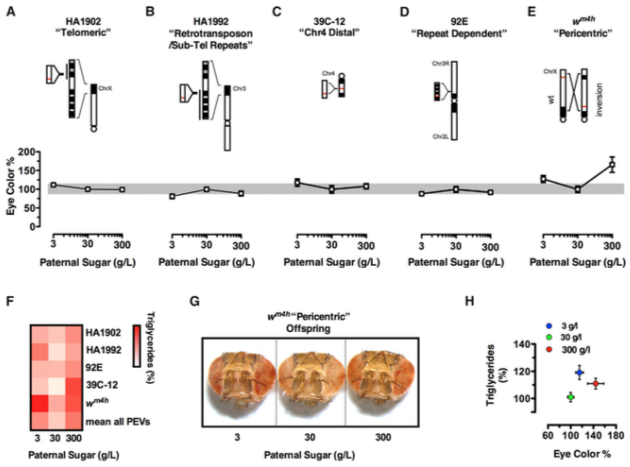


Figure 3. High Paternal Sugar Is a Physiological Su(var)

(A–E) (Top) Schematics of (A) *pP{RS5}5-HA-1902*, (B) *pP{RS5}5-HA-1992*, (C) *39C-12*, (D) *T(2;3)V21ePlacW (92E)*, (E) *In(1)w^m4h* PEV reporters. (Bottom) Eye pigment absorption (A_{480}) from heads of adult males, normalized to offspring of paternal 30 g/l sugar.

(F) Heatmap of paternal IGMR offspring from the PEV lines in (A–E). Triglyceride/weight normalized to paternal 30 g/l sugar.

(G) Representative median eye-colored w^{m4h} fly heads from 3, 30, and 300 g/l sugar-sired offspring.

(H) Correlation of eye color (heads) and offspring triglyceride/weight (carcass rest) of w^{m4h} flies measured in the same individuals.

Results are mean ± SEM of n = 3–8 experiments each with multiple replicates.

for *lipid particle*, the electron transport chain complexes I, IV, and V, *glycolysis*, *TCA cycle*, and *fatty acid metabolism* were all upregulated. These changes are consistent with energetics of enhanced lipid storage (Figure S2B). Three downregulated clusters were also detected, including *cell cycle and mitosis*, *body patterning*, and intriguingly, a cluster of *chromatin regulation* pathways. Consistent with sensitivity of the pericentric w^{m4h} reporter to IGMR *chromosome*, "*centromeric region*" was ranked second in the chromatin cluster and "*chromatin silencing*" ranked third. Examination of genes annotated as PEV suppressing, also known as *Su(var)*'s, revealed a concerted ~10%–20% downregulation, including members of most well-documented silencing pathways (Figure S2C and Tables S2 and S3). Thus, the IGMR embryo is characterized by gene expression favoring primary energy metabolism over chromatin control.

We next compared our data with chromatin mapping data sets from the community. Filion et al. used Dam-ID to annotate five major chromatin types, three repressive (black, blue, and green) and two active (red and yellow) (Filion et al., 2010). When intersecting our IGMR embryo data with their chromatin state maps, strong enrichment was observed in high-sugar sired embryos for genes embedded in "black" lamin/H1-associated heterochromatin and "blue" polycomb-associated chromatin, and relative depletion was observed for those annotated as "yellow," or housekeeping-type chromatin (Figures 4F and 4G). These findings were verified using rank-order (Figure 4F) and differential expression analyses (Figure 4G). No global effect was observed on "red" or "green" chromatin embedded genes. Consistent with these global indications of chromatin state dependency, the 68 significantly upregulated genes were almost exclusively found in "black" or "blue" chromatin while the 10 significantly downregulated transcripts were randomly distributed (Figure 4H). These data identify high paternal sugar as a chromatin-state-selective physiological Su(var) and identify IGMR as chromatin state dependent.

Polycomb and Core Heterochromatin Machinery Mediate Paternal IGMR

To genetically validate chromatin state regulation as a mechanistic underpinning of our model, we began systematically testing IGMR potential in mutants known to modify w^{m4h} variegation. We started with *Su(var)3-9^06*, a homozygous dominant suppressor allele of the H3K9 histone methyltransferase Su(var)3-9.

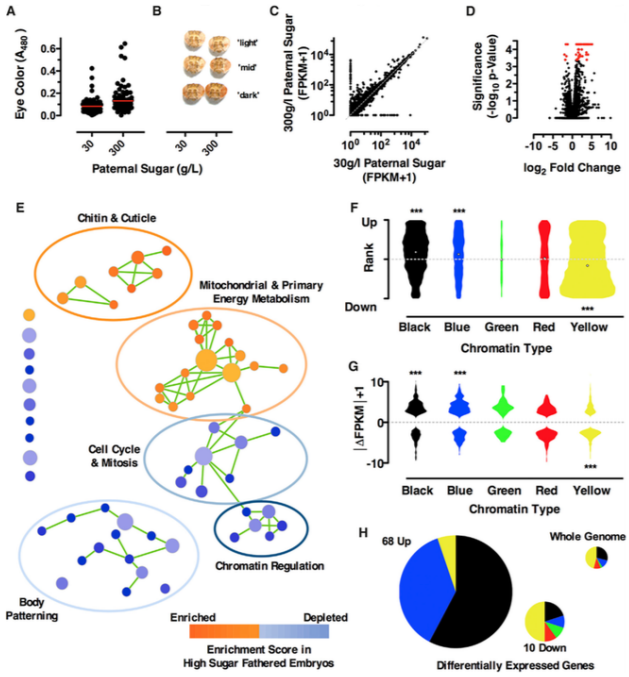


Figure 4. Paternal IGMR Alters Select Chromatin States in Offspring

(A) Interindividual variation of eye color of w^{1118} flies from fathers fed high (300 g/l) or medium (30 g/l) sugar food.

(B) Representative heads from w^{1118} offspring. Shown are heads representative of light, medium, and dark red eyes of each respective cohort.

(C–H) RNA-sequencing results of medium (30 g/l) and high-sugar (300 g/l) sired stage 17 embryos. (C) FPKM plot. (D) Volcano plot.

(E) Cytoscape enrichment map (p cutoff: 0.005, FDR Q-value cutoff: 0.025, overlap cutoff: 0.2) of gene set enrichment analysis (GSEA). (Orange) Gene sets enriched; (blue) gene sets depleted, in high-sugar IGMR. Color intensity reflects degree of enrichment. Major clusters are circled.

(F and G) (F) Rank and (G) absolute IGMR expression changes. Genes are allocated to one of five chromatin states (colors) according to their TSS (Filion et al., 2010). Plotted are (F) ranks for all genes and (G) absolute expression changes of the top 1,000 IGMR up and downregulated genes.

(H) Chromatin color annotation of all significantly up- and downregulated IGMR genes.

See also Figure S2.

Medium- and high-sugar-challenged *Su(var)3-9*^{06} fathers were mated with standardized w^{1118} females, and the resulting heterozygote offspring were monitored for adiposity (Figure 5A). Whereas w^{1118} animals reproducibly exhibited a ~10%–15% increase in adiposity upon high-sugar IGMR, F1 adult male offspring of *Su(var)3-9*^{06} fathers showed no intergenerational obesity response (Figure 5A). This provides genetic evidence that Su(var)3-9 is required for IGMR.

We also tested a second H3K9 methyltransferase, SetDB1. As heterozygotes, *SetDB1*^{1473} fathers gave both wild-type and mutant offspring. Intriguingly, both mutant (Figure 5B, red) and wild-type *SetDB1*^{1473} fathered offspring (Figure 5B, black)

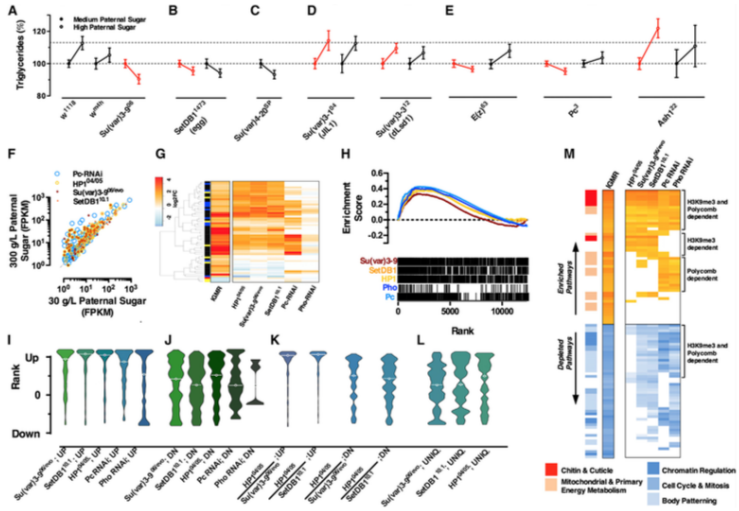


Figure 5. A Su(var)/PcG Axis Essential for Paternal IGMR

(A–E) Adiposity of offspring (triglycerides/weight) of mutant fathers challenged with medium (30 g/l; closed circles) or high sugar (300 g/l; open circles). Gray dashed line indicates normal w^{1118} IGMR response. IGMR adiposity responses are shown for offspring of (A) w^{1118}, w^{m4h}, and *Su(var)3-9*^{06}, (B) *SetDB1*^{1473}, (C) *Su(var)4-20*^{SP}, (D) *Su(var)3-1*^{04}, *Su(var)3-3*^{12}, (E) *Ez*$^{l(2)}$, *Pc*^{3}, and *Ash1*^{22} mutant (red) and wild-type (black) offspring. Results are mean ± SEM of n = 3–8 experiments each with multiple replicates.

(F) FPKM values of RNaseq data from medium and high-sugar-fathered embryos. 200 most upregulated genes from HP1 (open yellow), *Su(var)3-9* (closed cayenne), and *SetDB1* (closed orange) mutant first-instar larvae from Lundberg et al. (2013) and Pc-RNAi (open blue) experiment from Goodliffe et al (2007).

(G) Heatmap of expression changes of significantly changed genes in our paternal IGMR offspring embryo data set and in the Lundberg et al. *HP1*, *Su(var)3-9*, and *SetDB1* mutants and the Goodliffe et al (2007) *Pc*-, *Pho*-RNAi data sets.

(H) Enrichment plot for gene sets upregulated in *HP1*$^{04/05}$ (yellow), *Su(var)3-9*$^{06/evo}$ (cayenne), and *SetDB1*$^{10.1}$ (orange) mutants and *Pc*- (light blue) and *Pho*-RNAi (dark blue) in our stage 17 paternal IGMR offspring embryos.

(I–L) Violin plots of expression change distributions relative to all genes of stage 17 paternal IGMR offspring embryos for gene sets from *HP1*$^{04/05}$, *Su(var)3-9*$^{06/evo}$, and *SetDB1*$^{10.1}$ mutants (Lundberg et al.) and from *Pc* and *Pho*-RNAi embryos (Goodliffe et al.). IGMR relative rank is plotted for all available of the (I) 200 genes most upregulated and (J) 200 genes most downregulated by each mutant / RNAi line; (K) intersects of the 200 most up- or downregulated genes of the indicated pairs of mutants; and (L) genes in the 200 most upregulated gene sets unique to each respective mutant.

(M) Heatmap comparison of GSEA results from mutant and IGMR data. Plotted are the 50 most up- and downregulated pathways from paternal IGMR and respective scores from the mutant data sets. Colored bars left of the heatmap indicate clusters in Figure 4E.

completely failed to mount an IGMR obesity response (Figure 5B). *Drosophila* sperm develop as a syncytium, and therefore both mutant and wild-type sperm in such a cross will share a *SetDB1*^{1473} mutant cytosolic compartment for most of their development. These findings therefore indicate that SetDB1 in the male germline is necessary for proper IGMR.

H4K20me3 deposition follows H3K9me3 in the establishment of heterochromatin (Schotta et al., 2004). We therefore also tested *Su(var)4-20*^{SP}, a mutant for the H4K20 methyltransferase Su(var)4-20. As *Su(var)4-20*^{SP} is on Chr X, all male offspring from our crosses are wild-type. Again though, wild-type offspring will reflect the mutant heterozygosity of spermatogenesis. *Su(var)4-20*^{SP} fathers failed to transmit paternal IGMR to the F1 (Figure 5C). Thus, uncompromised expression of *Su(var)3-9*, *SetDB1*, and *Su(var)4-20* are absolutely required for IGMR. Of note, not all w^{m4h} suppressor alleles were IGMR incompetent. *Su(var)3-1*^{04} and *Su(var)3-3*^{12}, also known as Jil1 kinase and dLSD1, respectively, generated completely normal IGMR obesity responses (Figure 5D), thus indicating that IGMR is not directly linked to the w^{m4h} insertion locus itself. These findings

identify one of the first gene networks known to be absolutely required for proper intergenerational metabolic reprogramming.

Given the observed derepression in blue embedded genes (polycomb-associated; Figure 4), we tested IGMR potential in polycomb and trithorax group mutants. We found that, although *Ash1*^{22} mutants were fully IGMR competent, *Enhancer of zeste*, *E(z)*^{63} and *Polycomb*, *Pc*^{3} mutant males completely failed to elicit a response in the next generation (Figure 5E). Thus, polycomb- and H3K9me3-centric chromatin regulators are absolutely required for paternal diet-induced intergenerational obesity.

The IGMR Program Is Chromatin Encoded

To corroborate these findings, we compared our embryonic IGMR RNA-seq data with profiles from H3K9me3- and polycomb-insufficient mutants. We examined profiles from *Su(var)3-9*$^{06/evo}$, *SetDB1*$^{10.1}$, and *HP1*$^{04/05}$ mutant first-instar larvae (Lundberg et al., 2013) and *Pc*- and *Pho*-RNAi knockdown embryos (Goodliffe et al., 2007). Intriguingly, ~70% overlap was observed between our significantly dysregulated IGMR genes and those responsive to H3K9-centric or polycomb insufficiency (Figures 5F and 5G). The converse was equally true; each of the top 200 *Su(var)3-9*$^{06/evo}$, *SetDB1*$^{10.1}$, *HP1*$^{04/05}$, *Pho-RNAi*, and *Pc-RNAi* dysregulated gene sets showed strong enrichment in our high-sugar-sired F1 embryos (Figure 5H). Subgrouping confirmed specificity of these signals. First, transcripts upregulated by *Su(var)3-9*$^{06/evo}$, *SetDB1*$^{10.1}$, *HP1*$^{04/05}$, *Pho*, and *Pc* insufficiency (likely direct targets) showed clear coordinate increases in expression (Figure 5I) compared to apparently randomly distributed signals for transcripts downregulated by mutation (Figure 5J). Transcripts upregulated by both *HP1*$^{04/05}$ and either *Su(var)3-9*$^{06/evo}$ or *SetDB1*$^{10.1}$ (Figure 5K) showed much stronger signatures than transcripts significantly regulated by any one Su(var) mutant alone (Figure 5L). Thus, paternal IGMR mimics H3K9me3- and polycomb-dependent transcriptional dysregulation.

To test whether these signatures might directly contribute to metabolic reprogramming, we performed GSEA analysis of the Lundberg et al. (2013) and Goodliffe et al. (2007) data sets. Coordinate overlapping enrichment signatures were observed for key pathways of all five major IGMR clusters (Figure 5M and Table S2), including most chromatin and primary energy modules. Of note, the most significantly enriched pathways in our data set were those regulated by both silencing systems together (Figure 5M). Thus, IGMR is characterized by H3K9me3-/PcG-dependent dysregulation.

Sperm and Zygote Chromatin Plasticity Define IGMR

To gain further insight into IGMR transmission, we performed RNA sequencing from manually dissected and purified mature sperm of high- and medium-sugar-fed *w*^{1118} males (Table S4). Intriguingly, we again observed clear evidence of (1) broad transcriptional derepression in sperm of high-sugar-fed males (Figures 6A and 6B), (2) selective upregulation of black chromatin-embedded genes (Figure 6C), and (3) upregulation of *Su(var)3-9*$^{06/evo}$-sensitive genes (Figure 6D). These data indicate that transcriptional dysregulation in mature IGMR sperm is also chromatin state defined. In contrast to the embryo data, blue and yellow embedded genes appeared largely unaffected in the sperm transcriptome. Thus, chromatin-dependent signatures of IGMR are forecast in the P0 paternal germline.

Dysregulation of black embedded genes in both the sperm and zygote suggested potentially overlapping mechanisms for generation of the intergenerational signal in the germline and for hardwiring the IGMR phenotype in the offspring. To probe this idea genetically, we compared the effect of maternal versus paternal mutant allele contribution on IGMR. As described above, offspring of Su(var) and Polycomb mutant fathers were incapable of mounting an IGMR response (Figures 5A–5E and 6E, top row). In crosses in which *Su(var)3-9*$^{06/evo}$, *SetDB1*$^{10.1}$, or *Su(var)4-20*^{SF} mutations were contributed by the oocyte, IGMR-competent wild-type sperm were no longer able to evoke an intergenerational response (Figure 6E, bottom row). *Su(var)3-1*^{04} and *Su(var)3-3*^{12} mutants, unremarkable in the male germline, completely abrogated the response when contributed maternally (Figure 6E, bottom row). In contrast, oocytes contributing *Pc*^{3} and *E(z)*^{63} mutations, whose constitutive heterochromatin would not be predicted to be directly perturbed, mounted completely normal IGMR responses. Collectively, these data support a model in which IGMR results from and requires a permissive range of heterochromatin plasticity in the zygote.

To validate the idea, we intersected our embryo RNA-seq data with modENCODE H3K9me3 and H3K27me3 ChIP-seq profiles from same-stage embryos (16–20 hr) and from those isolated one time point earlier in development (12–16 hr), enabling us to gauge the dynamics of K9me3/K27me3 gain and loss (Nègre et al., 2011). We made several observations. First, IGMR-dysregulated genes represented a class undergoing highly dynamic H3K9 and H3K27 trimethylation (Figure 6F). This was true for our significantly changed IGMR genes, as well as the leading edge H3K9me3- and polycomb-dependent IGMR gene sets from Figure 5I (Figure 6F and data not shown). The bodies of these genes in particular were unmarked in 12–16 hr embryos and exhibit strong H3K9me3 and H3K27me3 just 4 hr later. Importantly, we observed the same signature when analyzing leading-edge genes of metabolic pathways upregulated in our obese IGMR phenotype (Figure S3A). Thus, genes undergoing highly dynamic H3K9me3- and H3K27me3-dependent silencing are specifically targeted for IGMR derepression.

Because repressive marks correlate with the higher-order chromatin structure and *cis*-regulatory domain organization, we also examined our gene sets in the context of insulator occupancy (Nègre et al., 2010). Analysis revealed that all three IGMR-dysregulated gene sets were on average far from class I (*CTCF*, *CP190*, and *BEAF*-associated) and were somewhat closer to class II (*SuHw*-associated) insulators (Figure 6G). These signatures were specific when compared to similarly expressed genes or to the entire transcriptome (Figure 6H, left). Intriguingly, the same signature was again evident in our most up- and downregulated sperm transcripts (Figure 6H, right). Thus, IGMR impacts spatially and chromatin-context-defined transcriptional units in fathers and in offspring.

Collectively, our data suggest that IGMR results from global alterations in chromatin state integrity within a permissive window, where obesity susceptibility results from reduced stage-specific epigenetic regulation of H3K27me3- and H3K9me3-defined domains. Our observations of *w*^{m4h} eye color desilencing (Figures 3

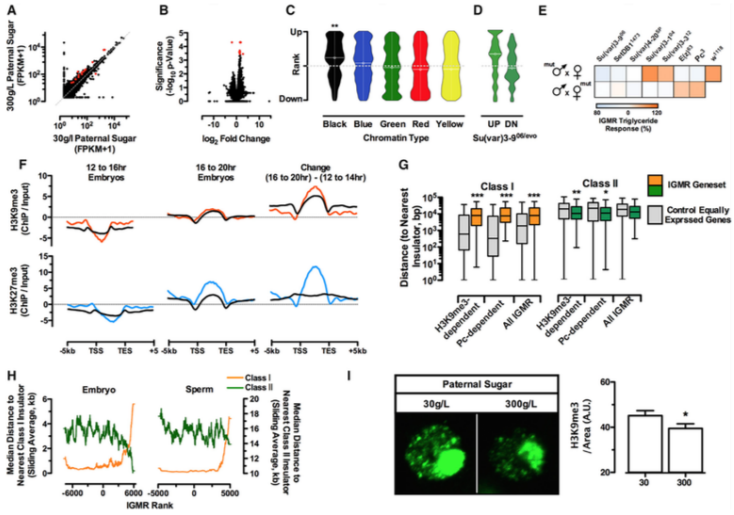


Figure 6. IGMR Signatures Are Forecast in the P0 Germline, and IGMR Changed Genes Show K9/K27me3 Dynamic Context

(A–D) RNA-sequencing results of sperm from medium- (30 g/l) and high-sugar (300 g/l) fed fathers; significantly changed genes are depicted in red. (A) FPKM plot. (B) Volcano plot. (C and D) IGMR expression changes in sperm of high- relative to medium-sugar-fed fathers (C) for the five chromatin colors according to Filion et al. (2010) and (D) for 200 most up- or downregulated genes from *Su(var)3-9*$^{06/evo}$ mutants from Lundberg et al (2013).

(E) Relative adiposity of male offspring (triglycerides/weight) (top row) from crosses of mutant fathers challenged with medium- or high-sugar diet, with w^{1118} mothers and (bottom row) of crosses of w^{1118} fathers with mutant mothers. The normal w^{1118} IGMR response is also shown (top row). Results are mean ± SEM of n = 3–8 experiments each with multiple replicates.

(F) ChIP/input signal from modENCODE data sets for leading-edge H3K9me3- and polycomb-dependent genes (red in top panels, blue in bottom panels) in our IGMR offspring embryo RNA-seq. H3K9me3 (top) and H3K27me3 (bottom) enrichment of 12- to 16-hr-old and 16- to 20-hr-old embryos (left) and the difference between the two stages (right). Black lines present the average for all genes.

(G) Box plots of distance to nearest class I and class II insulators. Shown are distances for leading-edge *Su(var)* and *PcG* upregulated genes in our IGMR offspring embryo RNA-seq. Grey boxes represent a control set of equally expressed genes. The boxes indicate the first and third quartiles, and the central line indicates the median. Whiskers extend to the most extreme data point, which is no more than 1.5 times the quartile range.

(H) Distance to nearest class I (orange) and class II (green) insulator plotted according to ranked expression change from IGMR RNA-seq results (high versus medium sugar). Values are sliding window averages of 500 genes.

(I) (Left) H3K9me3 staining of fat body cell nuclei from offspring of medium (30 g/l) and high-sugar (300 g/l) fed fathers. Results are mean ±SEM of n = 7 experiments, each with multiple replicates. (Right) Quantification of fat body cell nucleus H3K9me3 staining.

See also Figure S3.

and 4) and reductions in H3K9me3 immunofluorescence in adult IGMR offspring fat bodies (Figure 6I and Figure S3B) indicate that this chromatin state reprogramming is stable lifelong in the offspring.

A Conserved Signature for Chromatin-State-Associated Phenotypic Variation

More fundamentally, the above data identify a mechanism that directionally controls phenotypic variation within a population. To probe potential conservation of such processes, we searched for similar signatures in data sets from mouse and man. We examined two murine and three human microarray data sets focusing on adipose tissue from lean and obese individuals, first defining mouse and human ortholog pathways to all Flybase-annotated *Drosophila* Su(var)'s and then by using GSEA to test for dysregulation (Figures 7A–7D). Intriguingly, we observed clear signatures of Su(var) depletion in obese individuals in two of the most highly genetically controlled human adiposity data

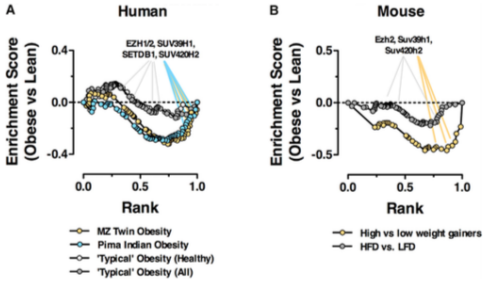


Figure 7. An IGMR Signature Conserved from Fly to Man

GSEA of mouse and human ortholog pathways for all Flybase-annotated *Drosophila* Su(var)'s (A) in three human adiposity data sets: 19 obese versus 20 non-obese Pima Indians by Lee et al. (2005); 13 human monozygotic (MZ) twin pairs, each discordant for obesity by Pietiläinen et al. (2008); and a human cohort for "typical" obesity by Klimcáková et al. (2011) and (B) in two murine obesity samples: surgically isolated adipose tissue of future high and low weight gainers from C57BL6/J mice biopsied prior to treatment with high-fat diet by Kozak et al. and diet-induced obesity comparing high- versus low-fat-diet-treated C57BL6/J animals by Voigt et al (2013).

sets available: first in a study of 19 obese versus 20 non-obese Pima Indians (Lee et al., 2005) and then, even more compelling, in a collection of 13 monozygotic twin pairs, each with one normal and one obese co-twin (Pietiläinen et al., 2008) (Figure 7A). Examination of the first figure in the latter study reveals clear evidence also of transcriptome-wide desilencing, with ~5-fold more up- versus downregulated genes in the obese co-twins. Further, similar signatures appear to predict murine obesity susceptibility. In an elegant study, Koza et al. isolated adipose tissue from young C57BL6/J mice prior to treatment with high-fat diet (Koza et al., 2006). Profiling the pretreatment samples from the lowest and highest weight gainers of the 107 animal strong cohort, the authors were able to establish predictive signatures for obesity susceptibility. Reanalyzing these data, we found clear evidence that Su(var) pathway depletion predicts obesity susceptibility (Figure 7B).

Leading-edge analysis of all three data sets revealed orthologs of our IGMR defining Su(var)3-9, Setdb1, Su(var)4-20, and E(z) regulators as driving the GSEA signal (Figures 7A and 7B, highlighted genes). Importantly, no obvious signatures were observed in an independent "typical" human obesity cohort in which the obesity is most likely driven by assorted genetic factors (Figure 7A) (Klimcáková et al., 2011), nor were they observed in diet-induced obese C57BL6/J mice (Figure 7B) (Voigt et al., 2013). Thus, Su(var) suppression characterizes obesity susceptibility on defined human and mouse genetic backgrounds.

These data identify conserved gene signatures for epigenetically defined phenotypic variation from fly to mouse to man.

DISCUSSION

Intergenerational Control of Chromatin State and Obesity

Here, we show that acute dietary interventions, as short as 24 hr, have the capacity to modify F1 offspring phenotype via the male germline. We show that reprogramming occurs in response to dietary manipulations over a physiological range and that phenotypic outcomes require polycomb- and H3K9me3-centric plasticity in spatially and chromatin-state-defined regions of the genome. The eye color shifts in w^{m4h} offspring (Figures 2E, 2G, and 2H) and the reduced fat body H3K9me3 staining in adult IGMR offspring (Figure 6I) supports the conclusions, first, that there are chromatin state changes and, second, that these are stable lifelong. These data are corroborated by selective derepression of *Su(var)3-9*, *SETBD1*, *Su(var)4-20*, and polycomb-sensitive transcripts (Figures 5F–5M); chromatin-state-associated transcriptional rearrangements genome wide (Figures 4F and 6C); selective reprogramming of highly dynamic histone-mark-defined regions (Figure 6F); and the fact that IGMR itself is sensitive to a string of distinct H3K9me3-centric and polycomb mutants (Figures 5A–5E). Although nontrivial, ChIP-seq comparisons of repressive chromatin architecture in mature sperm and multiple defined offspring tissues will be important to establishing the ubiquitousness of these regulatory events and the nature of intergenerational signal itself. These data highlight how acutely sensitive intergenerational control can be to even normal physiological changes, and they identify some of the first genes absolutely required for transmission.

Paternal Diet Regulates Chromatin-Defined Genes in the Germline and Offspring

First categorized simply as heterochromatin versus euchromatin, multiple empirical models now divide the genome into 5 to 51 chromatin states, depending on the analysis (Filion et al., 2010; Kharchenko et al., 2011; Ernst and Kellis, 2010). We find that paternal high sugar increases gene expression preferentially of heterochromatic-embedded genes in embryos. Specifically, these genes are characterized by active deposition of H3K9me3 and H3K27me3, by long distance from class I insulators, and by sensitivity to fully intact expression of *Su(var)3-9*, *Su(var)4-20*, *SetDB1*, *Pc*, and *E(z)*. The data support a model where phenotype has been evolutionarily encoded directly into the chromatin state of relevant loci. Specifically, an abundance of genes important to both cytosolic and mitochondrial metabolism appear to be embedded into H3K9me3- and distinct polycomb-dependent control regions. Indeed, our own GO analysis of the five chromatin colors from Filion et. al. (2010) indicate a largely mutually exclusive picture, in which functional pathways

are not randomly distributed across chromatin states (data not shown). Our paternal IGMR data set revealed clear and strong overlaps with pathways of black (lamin-associated) and blue (polycomb) chromatin and included many key metabolic pathways, including *glycolysis*, *TCA cycle*, mitochondrial OxPhos, *chitin*, and *polysaccharide metabolism*, changes that could well prime the system for altered functionality given the appropriate stimulus. Indeed, our paternal IGMR phenotype is a susceptibility to diet-induced obesity and is most readily observable upon high-sugar diet challenge.

Chromatin state coding of functional gene sets would provide a simple mechanism for transgenerational environmental response capable of rewiring even the earliest events of zygotic genome activation. The idea is also consistent with parallel avenues of research already in the literature. rRNA genes, for instance, are not only sensitive to the same Su(var)'s but are also known to influence metabolic gene expression and growth (Paredes et al., 2011). Flies with fewer rRNA genes (rDNA) exhibit a phenotype called bobbed (bb), which results in smaller bristles, a reduced growth rate, and a thinner chitinous cuticle (Ritossa et al., 1966). These phenotypes are intriguingly similar to the top GSEA enrichment clusters that we observed for IGMR, namely cell cycle, body morphogenesis, chitin deposition, and metabolism. Interestingly, the very same pathways (chitin synthesis, TCA cycle, carbohydrate, and lipid metabolism) are regulated by nutritional status in third-instar larva (Teleman et al., 2008), suggesting that the paternal IGMR signal acts to prime offspring for metabolic challenge.

Our data support a *trans*-acting mechanism. In the w^{m4h} experiments, male offspring inherited their X chromosome and thus the reporter from their unchallenged mothers, i.e., the reporter allele never encounters the initial signal but is reproducibly reprogrammed. Further, the failure of *Su(var)4-20*^{SH} and *SetDB1*^{1473} mutants to elicit IGMR responses in their wild-type offspring indicate that wild-type haploid sperm carry the same insufficient reprogramming template as their syncytial mutant counterparts. *cis*- and *trans*-acting mechanisms are not mutually exclusive though. Signals transmitted via paternal chromosomes, though likely transmitted in *cis*, may be manifest via expression of paternal transcripts, which then act in *trans*. Paternal reductions of Su(var)3-9, SetDB1, and Hp1, for instance, would affect the maternal genome in *trans*.

One Genotype, Multiple Paternally Directed Phenotypes

Despite their genetic similarity, isogenic or congenic animals reared under controlled conditions exhibit measurable variation in essentially all phenotypes. Such variability in genome output is thought to arise largely from probabilistic or chance developmental events in early life (Burga et al., 2011) (review in Whitelaw et al., 2010). Here, we map a mechanism that couples acute paternal feeding and zygotic chromatin state integrity directly to phenotypic output of the next generation. We find that these same signatures predict obesity susceptibility in isogenic mouse and human obesity cohorts. Because acute circadian fluctuations in feeding are essentially constant over evolutionary timescales, they are the perfect mechanistic input upon which a system could evolve to *ensure* defined phenotypic variation within a given population.

EXPERIMENTAL PROCEDURES

Fly Husbandry

Fly stocks were maintained on standard diet at 25°C on a 2 week generation cycle, ensuring a constant ancestral larva density. w^{1118} flies were single-sibling inbred for ten generations and maintained at a fixed fly density for another ten generations before experimental start. Fly strains used: In(1)w^{m4h} (w^{1e4h}), *pP{RS5}5-HA-1902* (*HA-1902*), *pP{RS5}5-HA-1992* (*HA-1992*), and *pP{RS5}5-HA-1925* (*HA-1925*) and *T(2;3)V21ePlacW* (92E), *Su(var)3-9*^{06}, *Su(var)2-5*^{05}, *Su(var)3-1*^{04}, *SetDB1*^{1473}, *Su(var)3-3*^{12} from (Phalke et al., 2009). *39C-12* from Sarah Elgin, *Pc*^{3} and *E(z)*^{63} from Leonie Ringrose. PEV lines were single-sibling inbred for ten generations.

Standard diet: Agar 12 g/l, yeast 18 g/l, soy flour 10 g/l, yellow cornmeal 80 g/l, molasses 22 g/l, malt extract 80 g/l, Nipagin 24 g/l, propionic acid 6,25 ml/l. Paternal diet intervention: Agar 12 g/l, yeast 10 g/l, propionic acid 4,5 ml/l, soy flour 30 g/l and white sugar as indicated.

Phenotyping

Body weight of five 7- to 12-day-old males flies was measured on a microbalance. Wing area determinations were made using ImageJ. Triglycerides (GPO Trinder, Sigma) and glucose and trehalose (Sigma; GAGO-20) were measured on centrifuged cleared lysates from groups of five flies crushed and sonicated in 100 µl RIPA buffer or TB buffer with or without trehalase (Sigma; T8778-1UN). "Café" assay was performed according to standard procedures. CO_2 production was quantified using a modification of Kucherenko et al. (2011). Eye pigment (A_{480}) was measured in centrifuge-cleared sonicates of one or five fly heads in 20/100 µl RIPA buffer, respectively. Fat body cryosections were fixed for 10 min in 2% formaldehyde in PBS, washed four times for 5 min each in PBS followed by immunofluorescence staining using rabbit anti-H3K9me3 (1:1000, upstate 07-442) and anti-Rabbit Alexa Fluor 488 (1:500; Molecular Probes). Confocal microscopy (LSM 700, Zeiss) analysis used Volocity 5.5 software (Perkin Elmer).

Sperm Dissection

Sperm dissection was modified from Dorus et al. (2006). See additional details in Extended Experimental Prodedures.

RNA Seqencing

Trizol-purified RNA was treated with Ribo-Zero (Epicenter) and libraries prepared with a TruSeq stranded kit (Illumina). > 15 million reads per sample were mapped using TopHat v2.0.8, with -G option against the *Drosophila melanogaster* genome (assembly BDGP5, Ensembl release 69). Gene expression values and significantly differentially expressed genes were calculated using Cuffdiff v2.1.1 with upper-quartile normalization and weighting multimapping reads (-N -u options).

Bioinformatic Analysis

Gene set enrichment analysis used GSEA 2.0 or GSEAPreranked with default parameters. Enrichment plots used the Cytoscape plugin Enrichment Map. Analysis of the five chromatin colors used BedTools (2.16.2). For microarray analyses, normalized probe values from the authors were mapped using Ensembl Biomart, and differential analysis against corresponding wild-types were performed using limma in R. Statistically significant was adjusted p value < 0.05 and fold change > 2. Enrichment of chromatin and insulator ChIP-seq data sets from modENCODE used deepTools 1.5.8.1 (Ramírez et al., 2014). Equivalently expressed gene sets were considered as the mean signal of the two genes ranked above and below each gene of interest. Distance to insulators was calculated using BedTools (2.16.2).

Statistical Analysis

Results are presented as means ± SEM. Statistical tests were performed using one-way ANOVA with a Newman-Keuls posttest. Statistical analysis of chromatin color data sets was a chi-square two-tailed analysis. All statistical analysis was done in with GraphPad Prism, unless otherwise noted.

ACCESSION NUMBERS

The RNA-sequencing data sets reported in this article have been deposited in the NCBI Gene Expression Omnibus and are accessible through GEO series accession number GSE62668.

SUPPLEMENTAL INFORMATION

Supplemental information includes Extended Experimental Procedures, three figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.11.005.

AUTHOR CONTRIBUTIONS

A.Ö. and J.A.P. conceived of the study. J.A.P., A.Ö., and A.L. designed the study, supervised all analyses, and wrote the manuscript. A.Ö., A.L., M.W., T.T., and M.D. did all fly work and analyzed the metabolic phenotypes. G.R., M.A., and N.I. contributed to study design and experiments. A.L., C.R., P.M.I., M.D., and A.Ô performed eating behavior and CO2 measurements. A.Ô. and A.L performed immunohistochemistry on fat body cells. A.Ô., A.L., M.W., T.T., U.B., M.S., N.R., and M.R. contributed to generation of the RNaseq data. E.C., S.H., T.V., and L.P. performed the bioinformatic analyses.

ACKNOWLEDGMENTS

This research was supported by the Max-Planck Society, EU (NoE Epigenesys), and the ERC 281641. A.Ô. was supported by Swedish VR K2011-78PK-21893-01-2 and SSMF grants. T.V., L.P., and E.C. were supported by Spanish Ministry grant BFU2011-30246, RYC-2010-07114, Marie Curie European Reintegration Grant "Evo-Chromo," and the IMPPC. C.R. and P.M.I. were supported by the Champalimaud Foundation, the Human Frontiers Program Project Grant RGP0022/2012, and the Portuguese Foundation for Science and Technology (FCT) grant PTDC/BIA-BCM/118684/2010. P.M.I. is supported by the postdoctoral fellowship SFRH/BPD/79325/2011 from the Foundation for Science and Technology. The authors are grateful to P. Georgiev, S. Raja, T. Manke, J. Longinotto, and T. Lu for important technical and theoretical help; to D. Corona, R. Bodmer, H. Esterbauer, and T. Jenuwein for critical discussions; and to H.J. and J.C.P. for everything.

Received: March 25, 2014
Revised: August 27, 2014
Accepted: October 31, 2014
Published: December 4, 2014

REFERENCES

Anderson, L.M., Riffle, L., Wilson, R., Travlos, G.S., Lubomirski, M.S., and Alvord, W.G. (2006). Preconceptional fasting of fathers alters serum glucose in offspring of mice. Nutrition *22*, 327–331.

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

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.

Blondeau, B., Avril, I., Duchene, B., and Bréant, B. (2002). Endocrine pancreas development is altered in foetuses from rats previously showing intra-uterine growth retardation in response to malnutrition. Diabetologia *45*, 394–401.

Braunschweig, M., Jagannathan, V., Gutzwiller, A., and Bee, G. (2012). Investigations on transgenerational epigenetic response down the male line in F2 pigs. PLoS ONE *7*, e30583.

Buescher, J.L., Musselman, L.P., Wilson, C.A., Lang, T., Keleher, M., Baranski, T.J., and Duncan, J.G. (2013). Evidence for transgenerational metabolic programming in Drosophila. Dis. Model. Mech. *6*, 1123–1132.

Burga, A., Casanueva, M.O., and Lehner, B. (2011). Predicting mutation outcome from early stochastic variation in genetic interaction partners. Nature *480*, 250–253.

Carone, B.R., Fauquier, L., Habib, N., Shea, J.M., Hart, C.E., Li, R., Bock, C., Li, C., Gu, H., Zamore, P.D., et al. (2010). Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. Cell *143*, 1084–1096.

Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat. Rev. Genet. *13*, 153–162.

Dias, B.G., and Ressler, K.J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. Nat. Neurosci. *17*, 89–96.

Dorus, S., Busby, S.A., Gerike, U., Shabanowitz, J., Hunt, D.F., and Karr, T.L. (2006). Genomic and functional evolution of the Drosophila melanogaster sperm proteome. Nat. Genet. *38*, 1440–1445.

Dunn, G.A., and Bale, T.L. (2009). Maternal high-fat diet promotes body length increases and insulin insensitivity in second-generation mice. Endocrinology *150*, 4999–5009.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat. Biotechnol. *28*, 817–825.

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., and van Steensel, B. (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212–224.

Fullston, T., Ohlsson Teague, E.M., Palmer, N.O., DeBlasio, M.J., Mitchell, M., Corbett, M., Print, C.G., Owens, J.A., and Lane, M. (2013). Paternal obesity initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content. FASEB J. *27*, 4226–4243.

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

Gniuli, D., Calcagno, A., Caristo, M.E., Mancuso, A., Macchi, V., Mingrone, G., and Vettor, R. (2008). Effects of high-fat diet exposure during fetal life on type 2 diabetes development in the progeny. J. Lipid Res. *49*, 1936–1945.

Goodliffe, J.M., Cole, M.D., and Wieschaus, E. (2007). Coordinated regulation of Myc trans-activation targets by Polycomb and the Trithorax group protein Ash1. BMC Mol. Biol. *8*, 40.

Greer, E.L., Maures, T.J., Ucar, D., Hauswirth, A.G., Mancini, E., Lim, J.P., Benayoun, B.A., Shi, Y., and Brunet, A. (2011). Transgenerational epigenetic inheritance of longevity in Caenorhabditis elegans. Nature *479*, 365–371.

Jimenez-Chillaron, J.C., Isganaitis, E., Charalambous, M., Gesta, S., Pentinat-Pelegrin, T., Faucette, R.R., Otis, J.P., Chow, A., Diaz, R., Ferguson-Smith, A., and Patti, M.E. (2009). Intergenerational transmission of glucose intolerance and obesity by in utero undernutrition in mice. Diabetes *58*, 460–468.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480–485.

Kiani, J., Grandjean, V., Liebers, R., Tuorto, F., Ghanbarian, H., Lyko, F., Cuzin, F., and Rassoulzadegan, M. (2013). RNA-mediated epigenetic heredity requires the cytosine methyltransferase Dnmt2. PLoS Genet. *9*, e1003498.

Klimcáková, E., Roussel, B., Márquez-Quiñones, A., Kovácová, Z., Kováciková, M., Combes, M., Siklová-Vítková, M., Hejnová, J., Srámková, P., Bouloumié, A., et al. (2011). Worsening of obesity and metabolic status yields similar molecular adaptations in human subcutaneous and visceral adipose tissue: decreased metabolism and increased immune response. J. Clin. Endocrinol. Metab. *96*, E73–E82.

Koza, R.A., Nikonova, L., Hogan, J., Rim, J.S., Mendoza, T., Faulk, C., Skaf, J., and Kozak, L.P. (2006). Changes in gene expression foreshadow diet-induced obesity in genetically identical mice. PLoS Genet. *2*, e81.

Kucherenko, M.M., Marrone, A.K., Rishko, V.M., Magliarelli, Hde.F., and Shcherbata, H.R. (2011). Stress and muscular dystrophy: a genetic screen for dystroglycan and dystrophin interactors in Drosophila identifies cellular stress response components. Dev. Biol. *352*, 228–242.

Lee, Y.H., Nair, S., Rousseau, E., Allison, D.B., Page, G.P., Tataranni, P.A., Bogardus, C., and Permana, P.A. (2005). Microarray profiling of isolated abdominal subcutaneous adipocytes from obese vs non-obese Pima Indians: increased expression of inflammation-related genes. Diabetologia *48*, 1776–1783.

Lee, H.C., Gu, W., Shirayama, M., Youngman, E., Conte, D., Jr., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell *150*, 78–87.

Lundberg, L.E., Stenberg, P., and Larsson, J. (2013). HP1a, Su(var)3-9, SETDB1 and POF stimulate or repress gene expression depending on genomic position, gene length and expression pattern in Drosophila melanogaster. Nucleic Acids Res. *41*, 4481–4494.

Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G.D. (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS ONE *5*, e13984.

Morgan, H.D., Sutherland, H.G., Martin, D.I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. Nat. Genet. *23*, 314–318.

Nègre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., et al. (2010). A comprehensive map of insulator elements for the Drosophila genome. PLoS Genet. *6*, e1000814.

Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., et al. (2011). A cis-regulatory map of the Drosophila genome. Nature *471*, 527–531.

Ng, S.F., Lin, R.C., Laybutt, D.R., Barres, R., Owens, J.A., and Morris, M.J. (2010). Chronic high-fat diet in fathers programs β-cell dysfunction in female rat offspring. Nature *467*, 963–966.

Ozanne, S.E., Wang, C.L., Dorling, M.W., and Petry, C.J. (1999). Dissection of the metabolic actions of insulin in adipocytes from early growth-retarded male rats. J. Endocrinol. *162*, 313–319.

Paredes, S., Branco, A.T., Hartl, D.L., Maggert, K.A., and Lemos, B. (2011). Ribosomal DNA deletions modulate genome-wide gene expression: "rDNA-sensitive" genes and natural variation. PLoS Genet. *7*, e1001376.

Patti, M.E. (2013). Intergenerational programming of metabolic disease: evidence from human populations and experimental animal models. Cell. Mol. Life Sci. *70*, 1597–1608.

Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M.C., and Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of Drosophila depends on the cytosine-5 methyltransferase DNMT2. Nat. Genet. *41*, 696–702.

Pietiläinen, K.H., Naukkarinen, J., Rissanen, A., Saharinen, J., Ellonen, P., Keränen, H., Suomalainen, A., Götz, A., Suortti, T., Yki-Järvinen, H., et al. (2008). Global transcript profiles of fat in monozygotic twins discordant for BMI: pathways behind acquired obesity. PLoS Med. *5*, e51.

Radford, E.J., Ito, M., Shi, H., Corish, J.A., Yamazawa, K., Isganaitis, E., Seisenberger, S., Hore, T.A., Reik, W., et al. (2014). In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. Science *345*. Published online July 10, 2014. http://dx.doi.org/10.1126/science.1255903.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42* (Web Server issue), W187–W91.

Rando, O.J. (2012). Daddy issues: paternal effects on phenotype. Cell *151*, 702–708.

Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature *441*, 469–474.

Rechavi, O., Houri-Ze'evi, L., Anava, S., Goh, W.S., Kerk, S.Y., Hannon, G.J., and Hobert, O. (2014). Starvation-induced transgenerational inheritance of small RNAs in C. elegans. Cell *158*, 277–287.

Ritossa, F.M., Atwood, K.C., and Spiegelman, S. (1966). A molecular explanation of the bobbed mutants of Drosophila as partial deficiencies of "ribosomal" DNA. Genetics *54*, 819–834.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. *18*, 1251–1262.

Seong, K.H., Li, D., Shimizu, H., Nakamura, R., and Ishii, S. (2011). Inheritance of stress-induced, ATF-2-dependent epigenetic change. Cell *145*, 1049–1061.

Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., Jr., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell *150*, 65–77.

Skorupa, D.A., Dervisefendic, A., Zwiener, J., and Pletcher, S.D. (2008). Dietary composition specifies consumption, obesity, and lifespan in Drosophila melanogaster. Aging Cell *7*, 478–490.

Teleman, A.A., Hietakangas, V., Sayadian, A.C., and Cohen, S.M. (2008). Nutritional control of protein biosynthetic capacity by insulin via Myc in Drosophila. Cell Metab. *7*, 21–32.

Voigt, A., Agnew, K., van Schothorst, E.M., Keijer, J., and Klaus, S. (2013). Short-term, high fat feeding-induced changes in white adipose tissue gene expression are highly predictive for long-term changes. Mol. Nutr. Food Res. *57*, 1423–1434.

Waterland, R.A., Dolinoy, D.C., Lin, J.R., Smith, C.A., Shi, X., and Tahiliani, K.G. (2006). Maternal methyl supplements increase offspring DNA methylation at Axin Fused. Genesis *44*, 401–406.

Whitelaw, N.C., Chong, S., and Whitelaw, E. (2010). Tuning in to noise: epigenetics and intangible variation. Dev. Cell *19*, 649–650.

Wolff, G.L., Kodell, R.L., Moore, S.R., and Cooney, C.A. (1998). Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J. *12*, 949–957.

Xing, Y., Shi, S., Le, L., Lee, C.A., Silver-Morse, L., and Li, W.X. (2007). Evidence for transgenerational transmission of epigenetic tumor susceptibility in Drosophila. PLoS Genet. *3*, 1598–1606.

REPORT

EPIGENETIC INHERITANCE

Transgenerational transmission of environmental information in *C. elegans*

Adam Klosin,[1,2] Eduard Casas,[3] Cristina Hidalgo-Carcedo,[1,2] Tanya Vavouri,[3,4]* Ben Lehner[1,2,5]*

The environment experienced by an animal can sometimes influence gene expression for one or a few subsequent generations. Here, we report the observation that a temperature-induced change in expression from a *Caenorhabditis elegans* heterochromatic gene array can endure for at least 14 generations. Inheritance is primarily in cis with the locus, occurs through both oocytes and sperm, and is associated with altered trimethylation of histone H3 lysine 9 (H3K9me3) before the onset of zygotic transcription. Expression profiling reveals that temperature-induced expression from endogenous repressed repeats can also be inherited for multiple generations. Long-lasting epigenetic memory of environmental change is therefore possible in this animal.

Resident animals are not the only ones subject to their environment; their progeny can also be affected (*1*–*11*). For example, starvation or exposure to high temperature in *Caenorhabditis elegans* can lead to altered small RNA transmission and putative target mRNA expression for up to three generations (*12*, *13*), and a few temperature-induced expression changes have been detected for two generations in animals with an inactive nuclear RNA interference (RNAi) pathway (*14*). In contrast, gene silencing initiated by exogenous double-stranded RNA (dsRNA) or piwi-interacting RNAs (piRNAs) can sometimes be stably inherited between generations (*15*–*19*).

When we subjected *C. elegans* to high temperature (25°C), expression from *daf-21 (Hsp90)* promoter::fluorescent protein constructs was strongly elevated (fig. S1). Expression from a single-copy transgene was still elevated in the progeny of animals transferred to 20°C after five generations at 25°C but not in their descendants (fig. S1A). In contrast, expression from an integrated multicopy array took 14 generations to return to basal levels after the temperature was reduced after 5 generations at 25°C (Fig. 1A and fig. S1). A single generation of growth at 25°C was sufficient to generate a seven-generation memory of increased expression (fig. S1C). Multigeneration inheritance of temperature-induced expression and a transgene-dependent phenotype was also observed with other high-copy arrays (table S1).

mRNA transcribed from a *daf-21* promoter array is first detected in wild-type (WT) worms at the 16-cell stage of development; this confirms no maternal supply of mRNA to the embryo (fig. S2) (*20*). Expression differences inherited from parents reared at different temperatures or sorted according to their expression were apparent from the onset of zygotic transcription (Fig. 1B and fig. S3), and genetic crosses demonstrated inheritance through both oocytes (Fig. 1C) and sperm (Fig. 1D). The array is therefore inherited in an inactive state but poised for a specific level of activation that reflects expression in the previous generation.

To distinguish whether inheritance occurs in cis with the DNA locus or in trans—for example, in the cytoplasm—we crossed worms with high and low expression to each other and then crossed the resulting F_1 male progeny to WT hermaphrodites (fig. S4) (*20*). The bimodal distribution of expression in the F_2 progeny indicates that the major mode of inheritance is in cis with the locus (Fig. 1E) (*21*).

To investigate chromatin modifications as potential mediators of this inheritance, we quantified histone modifications on the array in early embryos developing at 20°C whose grandparents had developed at either 16° or 25°C (Fig. 2A). Embryos whose grandparents developed at 25°C had less of the repressive histone modification H3K9me3 on the array than embryos whose grandparents developed at 16°C (Fig. 2, A and B). This difference was apparent in early embryos before the onset of zygotic transcription, indicating that the altered chromatin is not a secondary response to altered transcription in the embryo (Fig. 2, A and B). No differences were observed in the Polycomb-associated repressive modification trimethylated histone 3 lysine 27 (H3K27me3) or in H3K36me3 and H3K4me2, two modifications associated with active chromatin (Fig. 2B and fig. S5). The differences in H3K9me3 were maintained in late embryos after the onset of transcription (fig. S6).

No mRNA expression from the array was detected in the adult germ line (fig. S7). However, H3K9me3 was reduced on the array in the germline nuclei of adults that had been transferred from 16° to 25°C as embryos (Fig. 2, C and D, and fig. S8). Therefore, high temperature during germline development results in depletion of H3K9me3 from the array, even though there is no production of stable transcripts in this tissue.

The putative histone methyltransferase, SET-25, is responsible for all detectable H3K9me3 in *C. elegans* embryos (*22*) (fig. S5B), colocalizes with H3K9me3-enriched transgenic arrays within embryonic nuclei (*22*), and is required for the maintenance of piRNA-initiated stable gene silencing (*15*). Inactivating *set-25* increased expression from the array, with no difference in expression between animals maintained at 20° or 25°C (Fig. 3, A and B). Hence, the repression of the array at low temperature requires SET-25. Moreover, no differences in expression were observed between the F_1 offspring of *set-25* hermaphrodites mated with male animals transmitting an array with either high or low expression (Fig. 3B). In contrast, the inactivation of seven other small RNA pathway or chromatin components (including a Polycomb mutant *mes-2*) showed no obvious defects in the transmission of the expression memory (fig. S9). Even after >20 generations of growth at a constant temperature, substantial variation in transgene expression is observed in both WT and *set-25* mutant populations (Fig. 3C). In WT animals, these differences are transmitted to the next generation (Fig. 3C), but this is not the case in *set-25* mutants (Fig. 3C).

Our results suggest a simple model for how the transgene array shows memory of high-temperature exposure that endures for many generations (fig. S10). High temperature inhibits SET-25–mediated repression in the germ line, causing loss of H3K9me3 from the array. This derepressed chromatin is transmitted to subsequent generations, resulting in increased expression when transcription initiates in somatic lineages. Over multiple generations of growth at low temperature, repression is gradually restored by heterochromatin remodeling in each germline cycle. This is consistent with previously reported gradual quantitative intergenerational changes in H3K9me3 following a temperature change at some loci (*14*).

We tested whether this model predicts the behavior of endogenous loci in the genome by sequencing RNA from *set-25* mutants and WT animals at 20° and 25°C and from WT animals three generations after a change from 25° to 20°C. For protein-coding genes, derepression in *set-25* mutants provided weak prediction of increased

[1]EMBL-CRG Systems Biology Unit, Centre for Genomic Regulation (CRG), European Molecular Biology Organization, 08003 Barcelona, Spain. [2]Universitat Pompeu Fabra, 08003 Barcelona, Spain. [3]Program for Predictive and Personalized Medicine of Cancer, Institute Germans Trias I Pujol, Campus Can Ruti, 08916 Badalona, Barcelona, Spain. [4]Josep Carreras Leukaemia Research Institute, 08916 Badalona, Barcelona, Spain. [5]Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain.
*Corresponding author. Email: ben.lehner@crg.eu (B.L.); tvavouri@carrerasresearch.org (T.V.)

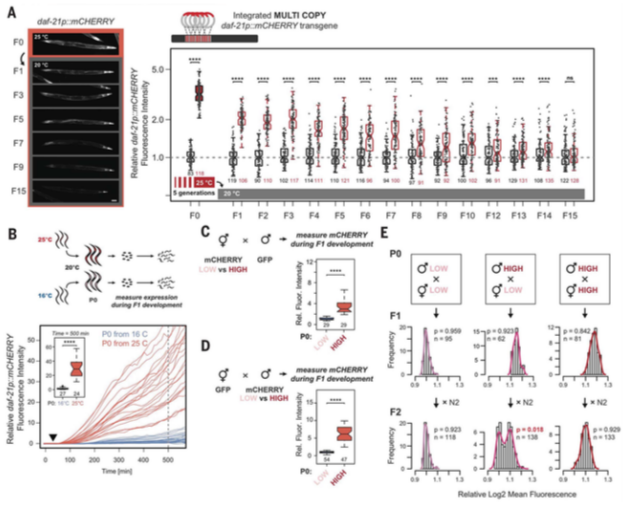


Fig. 1. Fourteen-generation memory of high temperature. (**A**) Adult expression of a *daf-21p::mCHERRY* integrated multicopy transgene at 20°C after five generations at 25°C. Scale bar, 0.1 mm. Stage-matched worms at 20°C are used as a reference for normalization (black). False-discovery rates (FDR) *q* values: ****$q < 0.0001$; ***$q < 0.001$; ns, $q > 0.05$ (Wilcoxon test). Sample size indicated. (**B**) Expression in embryos from animals transferred to 20°C at the L4 larval stage (inset: quantification at 500 min). Arrowhead indicates start of zygotic transcription of the transgene. Transmission occurs through oocytes (**C**) and sperm (**D**) and in cis with the locus (**E**). See fig. S4 for experimental design, intensities normalized to the "low" (low-expression) population; sample size and *P* value for Hartigans' dip test for unimodality. (B) (inset), (C), and (D), ****$P < 0.0001$.

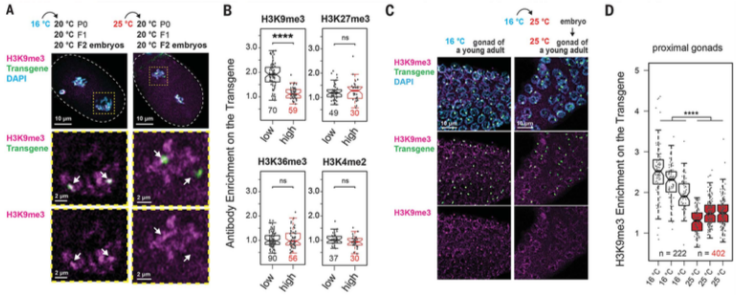
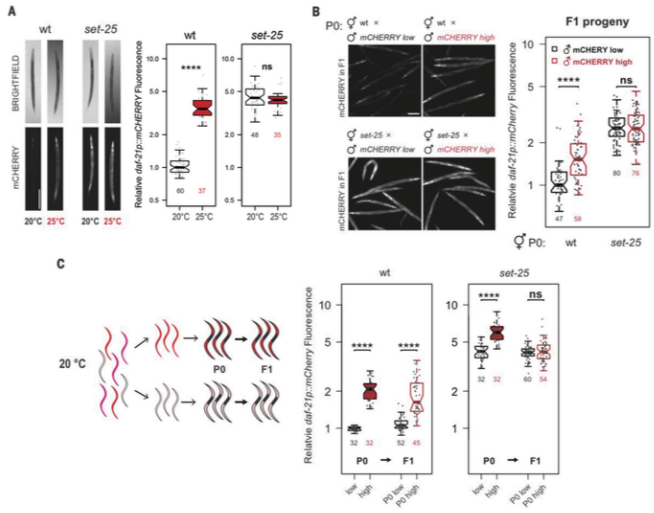


Fig. 2. Changes in H3K9me3. (**A** and **B**) H3K9me3 is depleted from the transgene locus in the F_2 descendants of animals grown at 25°C. L4 larvae from populations grown at 16° or 25°C were transferred to 20°C and cultivated until the following F_1 generation reached adulthood. The F_2 embryos were extracted and histone modifications quantified on the array by immunofluorescence combined with DNA fluorescence in situ hybridization (DNA FISH). (A) Representative two-cell–stage embryos stained with 4′,6-diamidino-2-phenylindole (DAPI) (blue), antibody against H3K9me3 (anti-H3K9me3) (pink), and a DNA FISH probe complementary to mCHERRY (green). Arrows indicate transgene loci. See fig. S5. (B) Quantification of histone modifications in early embryos. (**C** and **D**) Development at high temperature from embryo to adult results in reduced H3K9me3 on the array in the germline nuclei of adults. Gonads were extracted from adult worms shifted from 16° to 25°C during embryonic development, fixed, stained, and compared with those from animals kept constantly at 16°C. (C) Representative gonads stained with DAPI (blue), anti-H3K9me3 (pink), and a DNA FISH probe complementary to mCHERRY (green). (D) Each boxplot quantifies the nuclei of a single gonad (see also fig. S8). (B) and (D), ****$P < 0.0001$; (B) ns, not significant.



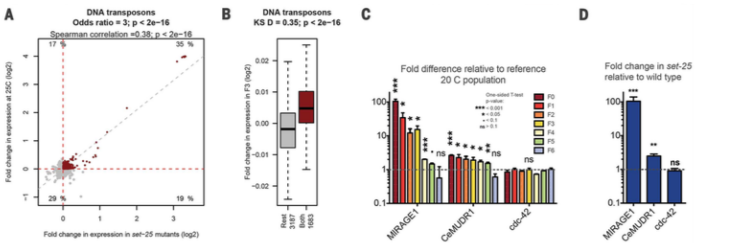


Fig. 3. Requirement for SET-25. (**A**) Quantification of *daf-21p::mCHERRY* expression in L4 larvae at 20° and 25°C in WT and *set-25* mutants. (**B**) Expression of a paternally derived transgene in the adult progeny of WT and *set-25* mutant mothers. A common batch of low- and (temperature-induced) high-expressing males was used. (**C**) Quantification of *daf-21p::mCHERRY* expression in the self-progeny (F_1) of parental (P0) animals sorted into high and low groups based on transgene expression at the L4 stage in WT and *set-25* mutants. ****$P < 0.0001$; ns, not significant. (A) and (B) Scale bars, 0.2 mm.

Fig. 4. Epigenetic expression memory of endogenous loci repressed by SET-25. (**A**) DNA transposon expression change in *set-25* mutants and at high temperature. Odds ratio quantifies the overlap (red loci, "both") between $\log_2$ fold change (FC) > 0. (**B**) FC expression three generations after a reduction in temperature from 25° to 20°C. Kolmogorov–Smirnov (KS) test statistic and *P* value are shown. See figs. S11 to S14 and table S4 for other repeats, protein coding genes, and analysis methods. (**C**) Expression of two DNA transposons at 25°C (F_0) and for six generations after decreasing the temperature to 20°C determined by quantitative PCR (table S3). *cdc-42* is a housekeeping gene as control. Expression is relative to animals grown at 20°C in parallel. (**D**) Expression of the same DNA transposons is increased in *set-25* mutants. **$P < 0.01$; ***$P < 0.001$; ns, not significant.

expression at high temperature (fig. S11), consistent with a larger contribution from other regulators, such as specific transcription factors. Derepression in *set-25* mutants was, however, a better predictor of increased expression at high temperature for multiple classes of repetitive elements and also for pseudogenes (Fig. 4A and figs. S11 to S14), consistent with impaired SET-25 activity's making an important contribution to the increased expression of many loci at high temperature.

Moreover, the increased expression of loci repressed by SET-25 with increased expression at high temperature was, although small, still detectable three generations after a return to low temperature (Fig. 4, A and B, and figs. S11 to

S14). Quantifying the expression of two DNA transposons by quantitative real-time polymerase chain reaction (PCR) in independent samples confirmed that their expression remained elevated for four and five generations after a return to 20°C (Fig. 4C and fig. S15). Their expression was also confirmed as SET-25–dependent (Fig. 4D).

Taken together, these results support the mechanistic model: At high temperature, SET-25 pathway activity is reduced, resulting in the derepression of many loci in the genome. After a return to low temperature, SET-25 activity is restored, but it takes multiple generations for repression to be completely reestablished. Expression from SET-25–repressed repeats therefore transmits information about a prior environmental exposure in this species.

In mammals, repressed repetitive elements can also escape epigenetic reprogramming (*23*, *24*) with variation in the expression of both individual repeats (*25*) and multicopy heterochromatic transgenes (*26*) being transmitted between generations. In flies, diet- (*6*) and stress-induced (*5*) changes in heterochromatin can also be transmitted for at least one generation. It is possible, therefore, that environmentally triggered changes in heterochromatin may provide a general mechanism for the epigenetic transmission of information between generations. It is interesting to speculate that the inheritance of environmentally triggered changes in expression from repressed chromatin may have been coopted to provide adaptive benefits to an organism.

REFERENCES AND NOTES

1. J. C. Jimenez-Chillaron *et al.*, *Diabetes* **58**, 460–468 (2009).
2. J. J. Remy, *Curr. Biol.* **20**, R877–R878 (2010).
3. S. F. Ng *et al.*, *Nature* **467**, 963–966 (2010).
4. B. R. Carone *et al.*, *Cell* **143**, 1084–1096 (2010).
5. K. H. Seong, D. Li, H. Shimizu, R. Nakamura, S. Ishii, *Cell* **145**, 1049–1061 (2011).
6. A. Öst *et al.*, *Cell* **159**, 1352–1364 (2014).
7. E. J. Radford *et al.*, *Science* **345**, 1255903 (2014).
8. D. Martínez *et al.*, *Cell Metab.* **19**, 941–951 (2014).
9. M. A. Jobson *et al.*, *Genetics* **201**, 201–212 (2015).
10. P. Huypens *et al.*, *Nat. Genet.* **48**, 497–499 (2016).
11. A. Klosin, B. Lehner, *Curr. Opin. Genet. Dev.* **36**, 41–49 (2016).
12. O. Rechavi *et al.*, *Cell* **158**, 277–287 (2014).
13. D. Schott, I. Yanai, C. P. Hunter, *Sci. Rep.* **4**, 7387 (2014).
14. J. Z. Ni *et al.*, *Epigenetics Chromatin* **9**, 3 (2016).
15. A. Ashe *et al.*, *Cell* **150**, 88–99 (2012).
16. M. J. Luteijn *et al.*, *EMBO J.* **31**, 3422–3430 (2012).
17. M. Shirayama *et al.*, *Cell* **150**, 65–77 (2012).
18. B. A. Buckley *et al.*, *Nature* **489**, 447–451 (2012).
19. S. G. Gu *et al.*, *Nat. Genet.* **44**, 157–164 (2012).
20. Materials and methods are available as supplementary materials.
21. S. Berry, M. Hartley, T. S. Olsson, C. Dean, M. Howard, *eLife* **4**, e07205 (2015).
22. B. D. Towbin *et al.*, *Cell* **150**, 934–947 (2012).
23. J. A. Hackett *et al.*, *Science* **339**, 448–452 (2013).
24. W. W. Tang *et al.*, *Cell* **161**, 1453–1467 (2015).
25. H. D. Morgan, H. G. Sutherland, D. I. K. Martin, E. Whitelaw, *Nat. Genet.* **23**, 314–318 (1999).
26. L. Daxinger, E. Whitelaw, *Nat. Rev. Genet.* **13**, 153–162 (2012).

ACKNOWLEDGMENTS

This work was supported by a European Research Council Consolidator grant (616434), the Spanish Ministry of Economy and Competitiveness (BFU2011-26206 and SEV-2012-0208), the AXA Research Fund, the Bettencourt Schueller Foundation, Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR), Framework Programme 7 project 4DCellFate (277899), and the European Molecular Biology Laboratory–Center for Genomic Regulation (CRG) Systems Biology Program. A.K. was partially supported by a la Caixa Fellowship. E.C. and T.V. were supported by the Spanish Ministry of Economy and Competitiveness (BFU2015-70581) and by an FI AGAUR Ph.D. fellowship to E.C. RNA-seq experiments were performed in the CRG Genomics core facility. RNA-seq data sets are deposited in the National Center for Biotechnology Information Gene Expression Omnibus under accession GSE83528.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/356/6335/320/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S4
References (27–35)

25 July 2016; resubmitted 21 December 2016
Accepted 24 March 2017
10.1126/science.aah6412

DEVELOPMENTAL GENETICS

Impaired DNA replication derepresses chromatin and generates a transgenerationally inherited epigenetic memory

Adam Klosin,[1,2]* Kadri Reis,[1,2]* Cristina Hidalgo-Carcedo,[1,2] Eduard Casas,[3] Tanya Vavouri,[3,4] Ben Lehner[1,2,5]†

Impaired DNA replication is a hallmark of cancer and a cause of genomic instability. We report that, in addition to causing genetic change, impaired DNA replication during embryonic development can have major epigenetic consequences for a genome. In a genome-wide screen, we identified impaired DNA replication as a cause of increased expression from a repressed transgene in *Caenorhabditis elegans*. The acquired expression state behaved as an "epiallele," being inherited for multiple generations before fully resetting. Derepression was not restricted to the transgene but was caused by a global reduction in heterochromatin-associated histone modifications due to the impaired retention of modified histones on DNA during replication in the early embryo. Impaired DNA replication during development can therefore globally derepress chromatin, creating new intergenerationally inherited epigenetic expression states.

INTRODUCTION

Multicopy transgene arrays are subject to epigenetic repression in the *Caenorhabditis elegans* germ line by the polycomb repressive complex 2 (PRC2) (*1*) and additional chromatin- and small RNA–related pathways (*2*–*4*). In *C. elegans*, modified histones and small RNAs are transmitted across generations (*5*, *6*), acting as carriers of epigenetic information (*7*–*10*). In addition to germline silencing, multicopy transgene arrays also show variation in their somatic expression level, which, at least in some cases, can be epigenetically inherited between generations (*11*, *12*).

RESULTS

To identify regulators of the heritable somatic repression of a *daf-21::mCherry* multicopy transgene array, we performed a genome-wide RNA interference (RNAi) screen (Fig. 1A). First-stage larval animals were fed in 96-well plates with bacteria expressing double-stranded RNA (dsRNA) targeting ~17,000 protein-coding genes, and expression from the array was scored in the adult worms of the same generation and in their larval progeny. Multiple RNAi clones that increased expression from the transgene targeted core components of the DNA replication machinery: DNA polymerase epsilon (*pole-1* or *pole-2*), the polymerase α-primase complex (*div-1*, *pri-2*, or Y47D3A.29/*POLA1*), replication factor C (*rfc-1* or *rfc-3*), and replication protein A (*rpa-2*) (Fig. 1B, fig. S1, and table S1).

The core replication machinery is mostly encoded by essential genes, but we could confirm the RNAi phenotypes using a hypomorphic allele, *or148*, of the gene encoding the B subunit of DNA polymerase α-primase, *div-1* (fig. S2) (*13*). This allele is a point mutation that causes delayed embryonic division due to prolonged S phase at 20°C and lethality at 25°C (*13*).

In *C. elegans*, the early stages of embryonic development are under maternal control (*14*). To test whether impaired DNA replication during embryonic development is sufficient to derepress the array, we crossed male animals carrying a *daf-21p::GFP* multicopy array to hermaphrodites carrying the *div-1* mutation (Fig. 1C). In this way, the array is delivered from a wild-type (wt) father into an egg produced by mutant *div-1* mothers, that is, containing mutant maternal *div-1* mRNA and DIV-1 protein. Expression in the resulting progeny was strongly up-regulated from the onset of zygotic transcription (Fig. 1, D to F). In contrast, crossing mutant *div-1* fathers to wt hermaphrodites carrying the array did not result in array derepression (fig. S3). Thus, impaired DNA replication during very early embryonic development results in increased expression from the start of zygotic transcription.

As in mammals, repressed chromatin in *C. elegans* is associated with specific histone modifications: trimethylation of histone H3 at lysine 27 (H3K27me3) and di- and trimethylation of H3 lysine 9 (H3K9me2/3) (*15*). In *C. elegans*, addition of H3K27me3 is catalyzed by the PRC2 (MES-2/3/6) complex (*16*). Inactivation of *mes-2* (Fig. 2, A and B) strongly increased expression from the transgene array. Similarly, inactivation of MET-2, a putative histone methyltransferase required for mono- and dimethylation of H3K9 (*3*, *17*), also strongly increased expression from the array (Fig. 2, A and B), as did inactivation of the putative H3K9me3 methyltransferase SET-25 (Fig. 2, A and B) (*3*). The very strong reduction in H3K9 methylation in a *met-2;set-25* double mutant (*3*, *18*) increased expression more than either single mutant (Fig. 2, A and B), and expression was highest in animals lacking H3K27me3 and H3K9me1/2/3 (*mes-2;met-2;set-25* triple mutants; Fig. 2, A and B), consistent with multiple repressive pathways being partially redundantly involved in repression of the array.

We tested the effects of impaired DNA replication in embryos lacking these histone modifications alone and in combination. Impaired replication still resulted in a strong increase in expression in animals lacking H3K27me3 (Fig. 2, C and D), indicating that the effects of impaired replication are not simply due to altered inclusion of this modification. Similarly, the array was still strongly up-regulated when replication was impaired in animals lacking all H3K9 methylation

[1]European Molecular Biology Laboratory–Centre for Genomic Regulation (EMBL-CRG) Systems Biology Unit, CRG, the Barcelona Institute of Science and Technology, 08003 Barcelona, Spain. [2]Universitat Pompeu Fabra, 08003 Barcelona, Spain. [3]Program of Predictive and Personalized Medicine of Cancer– Institute Germans Trias I Pujol, Can Ruti Campus, Ctra de Can Ruti, Camí de les Escoles s/n, Badalona, 08916 Barcelona, Spain. [4]Josep Carreras Leukaemia Research Institute (UC), Can Ruti Campus, Ctra de Can Ruti, Camí de les Escoles s/n, Badalona, 08916 Barcelona, Spain. [5]Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain.
*Co-first authors.
†Corresponding author. Email: ben.lehner@crg.eu

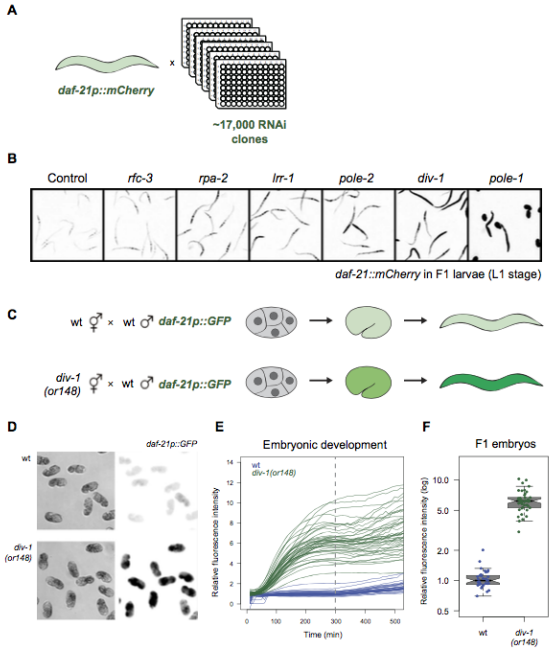


Fig. 1. Impaired DNA replication during embryonic development derepresses a transgene array. (**A**) Genome-wide RNAi screen to identify repressors of expression from a multicopy transgene array. (**B**) Expression of the *daf-21p::mCherry* transgene in F1 progeny when the indicated genes are inhibited by RNAi. (**C**) Male worms carrying a *daf-21p::GFP* multicopy transgene were crossed to wt or *div-1(or148)* mutant hermaphrodites. (**D** to **F**) Expression was quantified in F1 embryos by time-lapse microscopy. Quantification in (F) is at $t = 300$, indicated by the dashed line in (E) [6.2-fold difference, $P = 1.4 \times 10^{-46}$, two-sided t test; $n = 33$ and 42 for progeny of wt and *div-1* hermaphrodites, respectively]. Crossing male *div-1* animals to hermaphrodites carrying the *daf-21p::mCherry* did not result in an elevated transgene expression in the progeny (fig. S3), demonstrating that *div-1(or148)* heterozygosity in the progeny does not affect transgene expression during embryogenesis. Effects on additional transgenes are summarized in table S4.

(Fig. 2, C and D). Thus, the impact of impaired DNA replication cannot be due to alterations in just one of these repressive chromatin pathways. In contrast, the impact of impaired replication was strongly reduced in animals lacking both H3K27me3 and H3K9me1/2/3 (Fig. 2, C and D). This is not due to RNAi insensitivity or any saturation effect because inhibition of the chaperone HSP-1, which triggers a stress response and drives expression through the *daf-21* promoter, still strongly increased expression from the array (Fig. 2, C and D). Increased expression from the array after *pole-2(RNAi)* treatment was also partially suppressed in *mes-2;met-2;set-25* triple-mutant animals (fig. S4). This is consistent with impaired replication altering expression from the array by interfering with repression by multiple histone modifications (H3K27me3 and H3K9me1/2/3). In the absence of these modifications, impaired replication has a reduced effect on expression.

To characterize how the chromatin marks of the array are altered when replication is impaired, we first used chromatin immunoprecipitation (ChIP) to compare the levels of H3K27me3 in wt animals and in *div-1* mutants. H3K27me3 was reduced on the array in animals with impaired replication (~3- and ~4-fold in the gene body and the promoter, respectively) ($P < 0.01$) (fig. S5). Impaired replication therefore interferes with the maintenance of H3K27me3 on the array. However, H3K27me3 levels changed similarly on four additional regions of the genome (fig. S5), indicating that the alterations to chromatin are not restricted to the high-copy array.

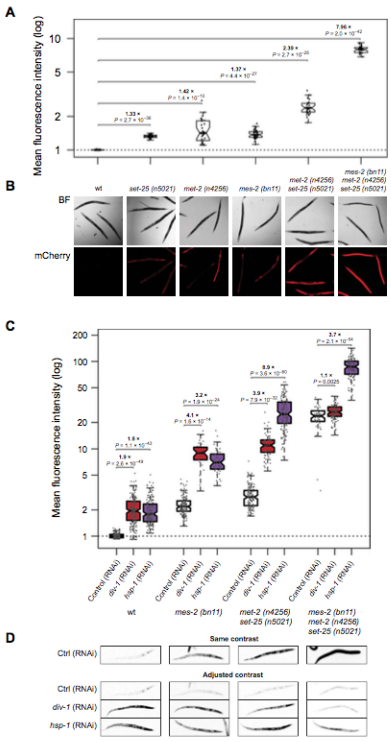


Fig. 2. The absence of repressive histone modifications suppresses the effect of impaired replication on expression. (**A**) Quantification of *daf-21p::mCherry* fluorescence intensity in adult worms with the indicated genotypes. Sample size: wt, 34; *set-25*, 37; *met-2*, 32; *mes-2*, 43; *met-2set-25*, 32; *mes-2;met-2;set-25*, 29. (**B**) Representative images of mCherry fluorescence. BF, bright field. (**C**) Quantification of *daf-21p::mCherry* fluorescence intensity in L1 larvae with the indicated genotypes when either *div-1* or *hsp-1* is inhibited by RNAi. Each dot represents one worm. The *y* axis is in log scale. *P* values were calculated by two-sided *t* test. Sample size: wt (control, 153; *div-1*, 203; *hsp-1*, 180), *mes-2* (control, 173; *div-1*, 32; *hsp-1*, 58), *met-2;set-25* (control, 145; *div-1*, 73; *hsp-1*, 153), *mes-2;met-2;set-25* (control, 55; *div-1*, 112; *hsp-1*, 97). (**D**) Representative images of mCherry fluorescence. Bottom: The contrast is adjusted for each genotype so that the change in expression relative to the control RNAi can be visualized. Ctrl, control.

To investigate this further, we quantified the global levels of H3K27me3 in wt and *div-1* chromatin from embryonic nuclei using immunofluorescence. Consistent with the ChIP results, H3K27me3 levels were globally reduced in the nuclei of early *div-1* mutant embryos (Fig. 3A; 1.35-fold, $P = 0.022$). We used the same technique to quantify the levels of H3K9me3 and found that they were also globally depleted in the chromatin of *div-1* embryos (Fig. 3B; 1.4-fold, $P = 0.025$). In contrast, the levels of a transcription activation–associated histone modification, H3K4me3, were increased in the *div-1* embryos (Fig. 3C; 1.4-fold, $P = 0.0055$). All the embryos quantified were from the 2- to 10-cell stage, which is before the onset of major zygotic transcription, indicating that the changes in chromatin are not a secondary consequence of any changes in transcription. We also observed similar changes in late-stage *div-1(RNAi)* embryos (fig. S6), indicating that the changes in chromatin are maintained after the onset of transcription and during development. Depleting *pole-2* confirmed the results obtained with *div-1(RNAi)* (fig. S7). The globally reduced levels of H3K27me3 and H3K9me3 were also confirmed by Western blotting (Fig. 3D and fig. S8). Impaired DNA replication therefore globally alters the histone modification levels in chromatin, including those before the onset of widespread zygotic transcription (*19*).

To test whether loci other than the transgene array also have altered expression when replication is impaired, we sequenced RNA from wt and *div-1* L1 stage larvae. Consistent with the response of the transgene array, many more genes had increased compared to decreased expression in the *div-1* mutants [493 up-regulated genes versus 9 down-regulated genes at a false discovery rate (FDR) of <0.05].

To relate changes in expression to the normal chromatin state of each gene, we used data from the modENCODE consortium (*15*). Consistent with the response of the array, this revealed widespread up-regulation of genes with normally repressed chromatin states (Fig. 4). This derepression was observed for genes normally characterized by states defined by either high H3K9me2/3 or high H3K27me3 (Fig. 4). In contrast, genes without repressive chromatin states (*15*) in wt animals were not up-regulated as a group in *div-1* mutants (Fig. 4). Together, these results show that impaired DNA replication during early development has a major impact on chromatin and gene expression, globally reducing the levels of repressive histone modifications and causing widespread up-regulation of heterochromatic genes in the resulting animals.

A recent study demonstrated that paternal histones marked with H3K27me3 are transmitted from *C. elegans* sperm chromatin to the zygote (*5*). These paternally inherited histones marked with H3K27me3 are then recycled during replication and deposited on the two daughter DNA strands during each of the early embryonic divisions, even in the absence of a functional PRC2 complex (*5*). We hypothesized that the loss of heterochromatin-associated histone marks might result from reduced retention of heterochromatic histones on the genome during the embryonic cell divisions. To test this, we quantified the decay of paternally inherited histones marked with H3K27me3 in control and *div-1 (RNAi)* embryos that were PRC2-deficient (Fig. 5). In this assay, there is no PRC2 activity in the early embryos because of the maternal *mes-2* genotype, and only the paternal genome contributes H3K27me3-marked histones to the zygote. The dilution of these modified histones during the early embryonic cell divisions therefore quantifies the extent to which they are successfully transmitted to and retained on the DNA daughter strands during each replication cycle. Compared to in control embryos, the decay of H3K27me3-modified histones during the early cell cycles was accelerated in *div-1(RNAi)* animals (Fig. 5, B and C). Thus, impaired DNA replication induces the loss of H3K27me3 by

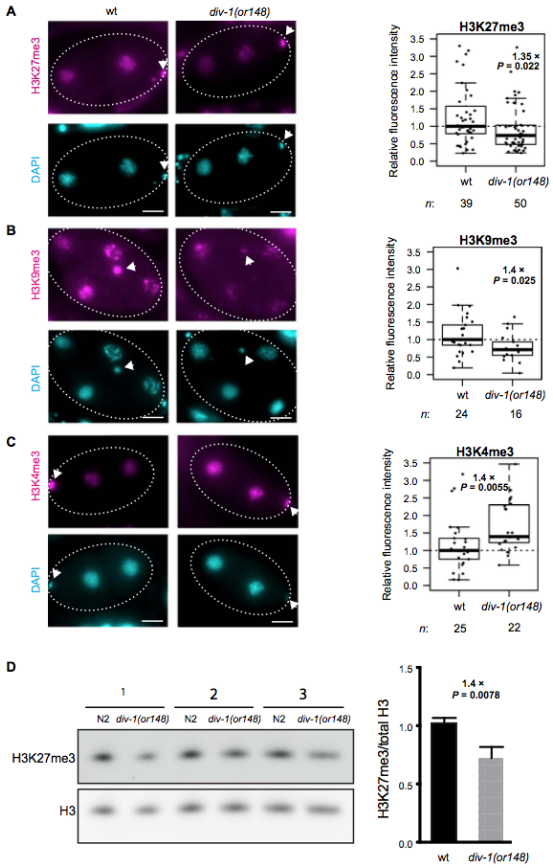


Fig. 3. Impaired DNA replication globally alters histone modifications. (**A** to **C**) Representative images and quantification of histone modification levels in wt and *div-1* embryos derived from self-fertilizing hermaphrodites. Average of each embryo after subtracting the background is plotted. Fold change relative to wt and *P* values (two-sided *t* test) is indicated for each comparison. DAPI, 4′,6-diamidino-2-phenylindole. Scale bars, 50 μm. (**D**) Western blot analysis showing H3K27me3 and total H3. Samples are three biological replicates from synchronized wt and *div-1(or148)* L1s from hermaphrodite parents. Quantification is shown on the right (means + SD; two-sided *t* test). Antibodies used here were validated by us and others (*5*, *12*).

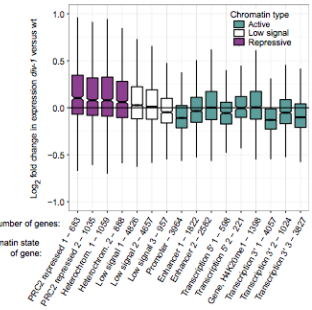


Fig. 4. Impaired DNA replication globally derepresses chromatin. Fold change in expression of genes mapping to different modENCODE chromatin states between *div-1* and wt L1 larvae. The number of genes assigned to each state is indicated.

impairing the retention of modified histones on the genome during the replication cycles of the early embryo. However, it is important to note that this assay does not exclude the possibility that impaired replication may also interfere with histone methyltransferase activity.

An important question in epigenetics is the extent to which acquired epigenetic states are transmitted between generations (*20*). We therefore tested what happens to the expression from the derepressed transgene array after normal DNA replication is restored. If the epigenetic state of the locus is not transmitted between generations, then restoration of normal DNA replication would result in the reestablishment of repression. In contrast, if the derepressed state is transmitted from parent to offspring, then expression would remain high in subsequent generations with normal replication.

To distinguish between these possibilities, we crossed wt males to *div-1(or148)* hermaphrodites (both carrying the *daf-21p::mCherry* transgene) and measured mCherry expression in the wt descendants for multiple generations (Fig. 6A). We found that the expression from the array was elevated for five generations after returning to the situation in which both animals and their parents had a wt *div-1* genotype (Fig. 6B). Moreover, introducing the *div-1(or148)* mutation for a single generation before outcrossing was sufficient to induce transgenerationally

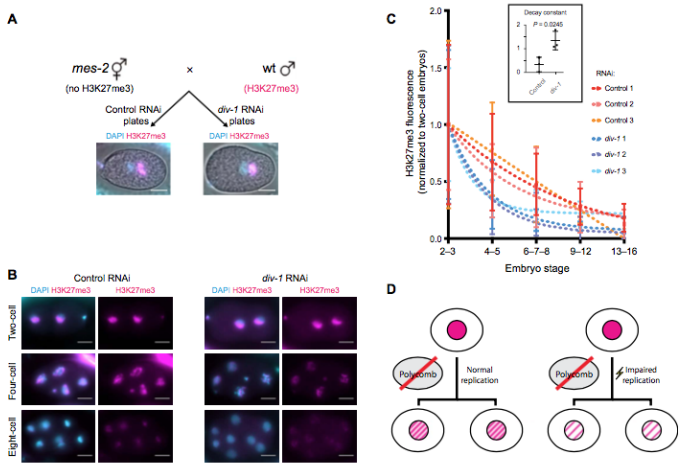


Fig. 5. Impaired replication interferes with the inheritance of H3K27me3-modified paternal histones. (**A**) wt males were crossed to *mes-2* mutant mothers either on control or *div-1(RNAi)* plates. Gravid worms were transferred to polylysine slides, gently squashed with a coverslip to allow embryos to extrude, and stained for H3K27me3. Images show H3K27me3 only in the paternal pronucleus. Scale bars, 50 μm. (**B**) Representative images of two-, four-, and eight-cell embryos from both groups stained with DAPI and an anti-H3K27me3 antibody. Scale bars, 50 μm. (**C**) Quantification of the H3K27me3 signal originating from the paternally deposited histones. Embryos were grouped according to the developmental stage, and a one-phase decay exponential curve was fitted through the data using Prism software. Decay rate constants for each curve are plotted in the inset. *P* values were calculated by two-sided *t* test. Sample size: control 1 (2 to 3, $n = 19$; 4 to 5, $n = 15$; 6 to 8, $n = 6$; 9 to 12, $n = 8$; 13 to 16, $n = 7$), control 2 (2 to 3, $n = 21$; 4 to 5, $n = 11$; 6 to 8, $n = 11$; 9 to 12, $n = 10$; 13 to 16, $n = 7$), control 3 (2 to 3, $n = 16$; 4 to 5, $n = 17$; 6 to 8, $n = 7$; 9 to 12, $n = 2$; 13 to 16, $n = 1$), *div-1* 1 (2 to 3, $n = 16$; 4 to 5, $n = 23$; 6 to 8, $n = 15$; 9 to 12, $n = 9$; 13 to 16, $n = 5$), *div-1* 2 (2 to 3, $n = 36$; 4 to 5, $n = 27$; 6 to 8, $n = 16$; 9 to 12, $n = 22$; 13 to 16, $n = 5$), *div-1* 3 (2 to 3, $n = 28$; 4 to 5, $n = 22$; 6 to 8, $n = 10$; 9 to 12, $n = 4$; 13 to 16, $n = 1$). Bars indicate SD, and the middle dots indicate the means. (**D**) Summary of the result of the experiment. Impaired DNA replication resulting from *div-1* knockdown interferes with efficient transmission of H3K27me3 histones to daughter nuclei.

inherited elevated expression (fig. S9). Thus, impaired DNA replication derepresses the transgene array, and this derepression takes multiple generations to completely reset after normal replication is restored. The return of the transgene expression to the basal level further demonstrates that the effect is epigenetic and not caused by genetic changes.

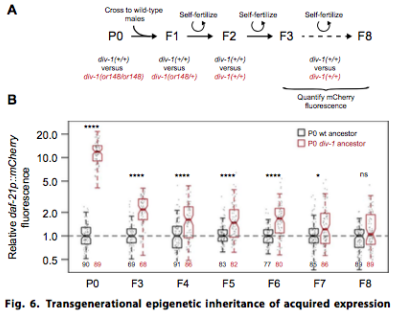


Fig. 6. Transgenerational epigenetic inheritance of acquired expression following DNA replication impairment. (**A**) Experimental design. (**B**) Quantification of *daf-21p::mCherry* fluorescence intensity in wt (black) and *div-1(or148)* (red) animals and in their wt descendants following outcrossing of the *or148* allele. At each generation, expression was normalized and compared to the control sample that was generated using the same batch of P0 males but crossed to wt ancestors and propagated in parallel in an analogous way. The *y* axis is in log scale. Sample size is indicated below each box plot. ****$P < 0.0001$; *$P < 0.05$; ns (not significant), $P > 0.05$ (Wilcoxon rank test).

DISCUSSION

Together, our results show that impaired DNA replication can have a major and directional epigenetic impact on a genome, resulting in a global loss of heterochromatic histone modifications, increased levels of euchromatic modifications, and increased expression from many normally heterochromatic genes. The mechanism underlying this is likely to be the impaired retention of modified histones on DNA during the early embryonic DNA replication cycles, as we have demonstrated for H3K27me3 (Fig. 7). Moreover, we have shown that the acquired expression changes following replication impairment can behave as "epialleles," persisting for multiple generations before resetting (Fig. 7). The contribution of the individual histone marks affected by the perturbed replication to the process of inheritance remains to be investigated. The observation that inhibiting several epigenetic regulators also generates multigeneration changes in gene expression and phenotypes in *C. elegans* suggests that perturbed chromatin states may frequently be transmitted between generations in this species (*8*, *12*, *21*). Transient overexpression of a histone demethylase during mouse sperm development leads to reduced survival and developmental abnormalities for three subsequent generations (*22*), suggesting that similar phenomena might occur in mammals.

Impaired replication is common in tumor cells (*23*) and also occurs during in vitro epigenetic reprogramming (*24*). Stalled replication forks can result in epigenetic alterations in chicken cells when repair pathways are inactivated (*25*), and persistent replication stress in *Drosophila* can generate polycomb-like phenotypes (*26*). In yeast, impaired DNA polymerase function (*27*) and deoxynucleotide triphosphate supply (*28*) can also have epigenetic consequences, and replication stress in mammalian cells alters the modifications detected on the pool of histones bound to the histone chaperone Asf1 (*29*). In future work, it will therefore be important to investigate the extent to which the rapid divisions and impaired cell cycle checkpoints (*30*–*32*) of early embryonic cells make their chromatin particularly vulnerable to impaired DNA replication, resulting

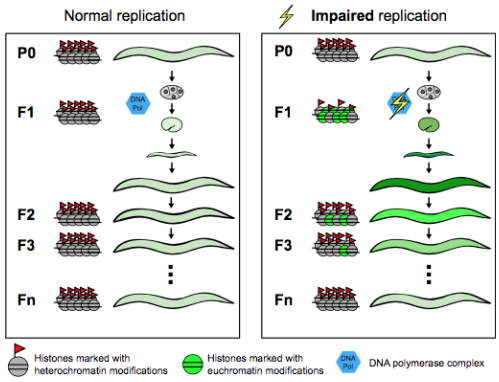


Fig. 7. Summary model. Impaired DNA replication during early embryonic divisions leads to inefficient retention of modified histones during chromatin replication, resulting in a reduction of heterochromatin-associated histone modifications and transgene derepression. Multiple generations are required to restore the repressed state.

in the global and directional changes in chromatin reported here. Moreover, it will be important to directly test the extent to which new epigenetic states in other species are transmitted through cell divisions and across generations, particularly given the limited transmission of modified histones to early embryos in mammals (*33*). Finally, we note that tumor cells also normally have impaired cell cycle checkpoints, suggesting that impaired DNA replication not only may be a driver of genomic instability but also perhaps may be a causal and directional driver of epigenetic alterations in cancer (*34*).

MATERIALS AND METHODS

Worm strains and culture conditions

All *C. elegans* strains used in this study are listed in table S2. Bristol N2 strain was used as the wt, and all other strains used were derived from it. Worms were cultured using standard conditions using NGM (nematode growth media) plates seeded with *Escherichia coli* OP-50 strain and grown at 20°C, including the *div-1(or148)* mutant strain, which is temperature-sensitive and exhibits 100% embryonic lethality at 25°C. The *daf-21p::mCherry* transgenic strain was generated by bombardment in an unc-119(ed3) background (*35*).

Genome-wide RNAi screen

The screen was carried out in high-throughput liquid feeding format in 96-well plates (*36*) using the Ahringer RNAi library (*37*). For the screen, a large number of embryos were harvested by bleaching, and the worms were hatched overnight in M9 to acquire a population of synchronized L1s. In the feeding plates, every well contained culture of one transformed *E. coli* clone, grown in 800 µl of LB + ampicillin (Amp) overnight at 37°C at 220 rpm. To set up the feeding, the worms were counted and diluted to have 5 worms/µl, and 10 µl of worms was dispensed into each well to have 50 to 75 worms per well. One hour before adding the *E. coli* cultures to the wells, double-stranded RNA synthesis was induced by adding 4 µl of 1 M isopropyl-β-D-thiogalactopyranoside (IPTG) to the cultures, which were then grown for 1 hour at 37°C at 220 rpm. The bacteria were pelleted at 2500*g* for 5 min and resuspended in 100 µl of NGM + Amp and IPTG. A total of 40 µl of the resuspended bacteria was added to each well of 50 to 75 worms. The worms were allowed to grow until most of the food had been consumed and the worms were gravid with some L1 larvae around (88 to 90 hours). Each 96-well plate included several wells of feeding with control RNAi to be used as a reference well for screening. The primary screen was carried out by eye with a Leica DMI6000 B microscope with a Lumen 200 metal arc lamp, observing the intensity of the transgene expression under the microscope with ×10 magnification. The secondary more stringent screen was carried out feeding the worms in triplicate with all the primary hits.

Additional RNAi experiments

The control RNAi strain used for all experiments produces dsRNA that does not map to any expressed sequence of the genome (Y95B8A_84.g). For the experiment in Fig. 2, the RNAi bacteria were spotted on NGM plates containing Amp and IPTG (*38*). Synchronized L4 worms were then transferred from OP50 plates and grown at 20°C for 24 hours. The F1 embryos were extracted from gravid animals through hypochloride treatment and allowed to hatch overnight in M9 buffer. Successfully hatched L1 animals were examined for expression of mCherry. For the *mes-2* and *met-2;mes-2;set-25* mutants, P0 worms showing a clear uncoordinated phenotype (hence homozygous for *mes-2* mutation) were picked to ensure the correct genotype of the mothers. *hsp-1* RNAi bacteria were diluted with three parts of control RNAi strain for one part of *hsp-1* RNAi to reduce the severity of the phenotype and allow examination of expression in F1 progeny.

Time-lapse microscopy and expression quantification

When the goal was to analyze the progeny after the cross, the crosses were carried out at 20°C by picking about one male per hermaphrodite (L4 larvae). After 24 hours, for each cross, 60 fertilized hermaphrodites were transferred to a well of concave glass slide containing 50 µl of phosphate-buffered saline (PBS). The embryos were released by cutting the worms with surgical needles. The embryos were washed briefly with an equal volume of 5% hypochlorite solution followed by four rounds of washing with 2 volumes of PBS. Four-cell stage were collected with a mouth pipette, washed twice in 200 µl of PBS and transferred to the 96-well plates (optical bottom; Nunc) with 100 µl of PBS. The embryos were carefully moved with an eyelash pick to the center of the field view and imaged with a Leica DMI6000 B microscope. The two biological samples were processed in parallel, each on a separate slide. The selection of staged embryos was done within 5 min for each sample to ensure the synchrony of the embryos. For time lapse, the images were taken with a 10× objective every 10 min for 16 hours in bright field and green (green fluorescent protein) channels. The images were analyzed with ImageJ where the embryos were first selected in bright field, and the selection was transferred to the fluorescent images, from where the level of expression was quantified. For each embryo, a background area was selected from close proximity to which the intensity ("integrated intensity" in imageJ) was then normalized. The same process was carried out for all the time points, giving a transgene expression intensity curve that was visualized with R (version 2.15.3). All subsequent analysis was carried out with R.

ChIP–quantitative polymerase chain reaction

Synchronized worms were grown on OP50 feeding plates to obtain gravid adults (about 65 to 70 hours after L1) that were collected in M9. The samples were fixed in 1.5% formaldehyde at room temperature for 30 min and quenched with 0.25 M glycine at room temperature for 15 min. The samples were washed twice with M9 and once with FA buffer [50 mM Hepes-KOH (pH 7.5), 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 150 mM NaCl] with protease inhibitors (Mini EDTA-free, Roche cOmplete). The worm pellet was snap-frozen, and FA buffer (+ protease inhibitors) was added to the samples and sonicated using the Bioruptor Sonication System Diagenode version 1.1, at 4°C for 20 cycles (high power, 30 s on + 30 s off) to achieve chromatin fragment sizes of 200 to 600 base pairs (bp) and was centrifuged for 25 min at 4°C at 13,000 rpm. Protein (0.1 mg) was added into the primary antibody reaction in FA buffer (+ protease inhibitors) up to 500 µl with 2 µl of H3K27me3 (Millipore 07-449) antibody. From each experiment, 1% of the volume was stored as an input control at −20°C before adding the antibody. The first antibody reaction was rotated at 4°C overnight. The next day, 30 µl of unblocked protein A beads (Diagenode catalog no. kch-503-880) was washed with 500 µl of FA buffer. The chromatin/first antibody mix was added to the beads, followed by 2-hour rotation at 4°C. The beads were washed thrice with 1 ml of low-salt buffer [50 mM Hepes-KOH (pH 7.5), 150 mM NaCl, 1% Triton X-100] and once with 1 ml of high-salt buffer [50 mM Hepes-KOH (pH 7.5), 500 mM NaCl, and 1% Triton X-100]. All centrifugations were carried out for 3 min at 4°C at 3000 rpm. After the last wash, the beads were left to dry, and the samples were eluted along with the input samples in 100 µl of fresh elution buffer

(1% SDS with 0.1 M $NaHCO_3$) for 3 hours at 65°C at 1100 rpm. The beads were centrifuged at 3000 rpm, the supernatant was purified with QIAquick PCR Purification Kit (#28104), and the samples were eluted in 200 µl of polymerase chain reaction (PCR)–grade water. Quantitative PCR (qPCR) reactions were run in LightCycler 480 Multiwell Plate 384, each well containing 2 µl of the sample, 5 µl of the 2× LightCycler 480 SYBR Green I Master Mix, and 1 µM of reverse and forward primers.

From the qPCR analysis, the cycle threshold was first normalized to the input and with the Δ-Δ method (*39*), and the percent input was calculated. For each antibody, this value was then normalized to the percent input value of H3 total histone. For each biological replicate, two technical replicates were analyzed.

Immunofluorescence

Immunofluorescence was performed using a method adapted from the approach of the Seydoux laboratory (*40*). Gravid worms were picked into 5 µl of M9 on top of a polylysine-coated slide. A coverslip was placed, and the worms were gently squashed to allow embryos to extrude. The embryos were immediately freeze-cracked on liquid nitrogen and fixed with methanol for 10 min, followed by acetone for an additional 10 min. After three washes in PBS containing 0.25% Triton X-100 (PBS-T), slides were blocked in PBS-T with 0.5% bovine serum albumin (BSA) before overnight incubation with primary antibody [H3K9me3 (#61013, Active Motif), H3K4me3 (#ab8580, Abcam), or H3K27me3 (#07-449, Millipore)] at 4°C. The slides were then washed with PBS-T and incubated for 2 hours with secondary antibody (Alexa 555, Invitrogen) at room temperature. After three washes in PBS-T, the samples were mounted in Fluoroshield with DAPI mounting medium (Sigma). Images were taken using a Leica DMI6000 B microscope. Quantification was performed using ImageJ. Chromatin masks for each nucleus were created using the DAPI channel. Using these masks, histone modification fluorescence for each embryo was measured as the average of all the nuclei in the embryo. For each embryo, a background area with no nuclei was selected to which the intensity was then normalized. Embryo average fluorescence after subtracting the background was plotted. Metaphase nuclei were not taken into account for quantification. Considering only interphase nuclei did not change these results (fig. S10).

Western blotting

Synchronized L1 animals were obtained by bleaching gravid adults, followed by three washes and an overnight hatch in M9. Worms were counted, resuspended in sample buffer [2.36% SDS, 9.43% glycerol, 5% β-mercaptoethanol, 0.0945 M tris-HCl (pH 6.8), 0.001% bromophenol blue], snap-frozen in liquid nitrogen, and sonicated three times for 30 s at 15 W. Samples were then boiled for 2 min and loaded on a precast SDS–polyacrylamide gel electrophoresis (15%) gel (Invitrogen). Transfer to nitrocellulose membranes was done using iBlot (Thermo Fisher) system. The membranes were incubated with primary antibodies for 1 hour [H3K27me3 (#07-449, Millipore), H3 (ab1791, Abcam), and H3K9me3 (ab8898, Abcam)] and washed six times with PBS-T + BSA, followed by an incubation with horseradish peroxidase–conjugated anti-rabbit secondary antibody. For Fig. 3, membranes were stripped after exposure to H3K27me3 antibody. Secondary antibody incubation after stripping confirmed that H3K27me3 antibody had been completely removed before probing for total H3 as loading control. Images were taken with Amersham Imager 600, and quantification was performed with ImageJ.

RNA sequencing

L1 wt and *div-1* larvae were harvested in triplicate, and RNA was extracted using TRIzol and freeze-cracking. Libraries were prepared using TruSeq Stranded mRNA Sample Prep Kit version 3 (reference no. RS-122-2101/2, Illumina) according to the manufacturer's protocol. Briefly, 1 µg of total RNA was used for poly(A)-mRNA selection using streptavidin-coated magnetic beads and was subsequently fragmented to approximately 300 bp. Complementary DNA (cDNA) was synthesized using reverse transcriptase (SuperScript II, reference no. 18064-014, Invitrogen) and random primers. The second strand of the cDNA incorporated deoxyuridine triphosphate in place of deoxythymidine triphosphate. Double-stranded DNA was further used for library preparation and was subjected to A-tailing and ligation of the barcoded TruSeq adapters. Library amplification was performed by PCR using the primer cocktail supplied in the kit. All purification steps were performed using AMPure XP Beads (reference no. A63880, Beckman Coulter). Final libraries were analyzed using an Agilent DNA 1000 chip to estimate the quantity and to check size distribution and were then quantified by qPCR using the KAPA Library Quantification Kit (reference no. KK4835, Kapa Biosystems) before amplification with Illumina's cBot. Indexed libraries were loaded at a concentration of 2 pM onto the flow cell (12 pM per lane) and were sequenced 1 × 50 on Illumina's HiSeq 2000.

Sequence reads were mapped using TopHat2 version 2.1.0 (*41*), with default parameters against a custom genome consisting of the *C. elegans* genome assembly WS215 from WormBase and the sequence of the transgene vector. Reads aligning to different genomic features were counted using featureCounts version 1.5.0 (*42*) with the option -s 2 -M --fraction to include multimapping reads and weighting them by number of matches. We used the *C. elegans* genome annotation from Ensembl Release 70. Data scaling, normalization, and tests for differential expression were performed using DESeq2 package version 1.8.1 (*43*) for R 3.2.0 (R Core Team 2015). Chromatin state segmentation and description were from modENCODE (*15*) using the early-stage embryo chromatin. Each gene was assigned to all overlapping states.

Intergenerational inheritance of expression changes

Male P0 worms homozygous for the *daf-21p::mCherry* array where crossed to P0 hermaphrodites carrying the *daf-21p::mCherry* array in addition to either a wt [denoted as *div-1(+)*] or mutant *div-1(or148)* allele. Expression was quantified in these P0 hermaphrodites. F1 hermaphrodite progeny were picked at the L4 stage to separate plates and allowed to self. After 3 days, single F2 progeny L4 hermaphrodites were transferred to separate wells. Two days later, the adult worms were removed (leaving multiple laid F3 embryos and larvae in the well) and genotyped using PCR (forward primer, gaacggagcacttgggaaga; reverse primer, tgttcgtgggaccaatgaga), followed by 1-hour restriction digest with Bsr GI (New England Biolabs), cutting only the *div-1(or148)* allele product. F3 progeny of F2 worms identified as wt were subsequently followed and analyzed for expression of mCherry for several generations. All worms were picked and handled at a standard dissecting microscope with no fluorescence to avoid any biases. At each generation, mean fluorescence intensity in day-1 adults was used as a readout. We normalized the expression of div-1 descendants to the median expression of descendants of wt worms propagated in parallel.

Statistical analyses

Unless otherwise stated, all statistical analyses were performed using R (version 2.15.3). Box plots in all figure plots indicate median and first

and third quartiles. Lower and upper whiskers extend to 1.5× the interquartile range (IQR) from the first and third quartiles, respectively. Notches, where present, extend $1.57 \times \mathrm{IQR}/\sqrt{n}$. Sample size depended on the ease of preparation and was larger for L1 animals that were treated, imaged, and analyzed in bulk, in contrast to adult animals that were picked manually under a dissecting scope before imaging, resulting in a relatively smaller sample size. No animals were excluded from the analysis. No blinding or randomization was used.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/3/8/e1701143/DC1

fig. S1. Expression of the *daf-21p::mCherry* reporter in the progeny of animals treated with RNAi targeting different subunits of the DNA polymerase complex and its associated proteins.
fig. S2. Increased transgene expression in *div-1* mutants.
fig. S3. Maternal *div-1* deficiency results in elevated transgene expression in the offspring.
fig. S4. Transgene up-regulation following *pole-2* knockdown is suppressed in the *mes-2;met-2; set-25* triple-mutant background.
fig. S5. Impaired DNA replication reduces H3K27me3 levels on multiple loci.
fig. S6. Global reduction of repressive histone marks and a gain of activating histone marks in late *div-1(or148)* embryos.
fig. S7. Knockdown of *pole-2* results in reduction of H3K27me3 mark and increase in H3K4me3 in early embryonic chromatin.
fig. S8. Reduction in H3K9me3 mark in *div-1(or148)* mutant L1s detected by Western blot.
fig. S9. Passage of the transgenic array through impaired replication for a single generation is sufficient to trigger a multigenerational effect.
fig. S10. Quantification of H3K27me3 in interphase nuclei.
table S1. List of genes whose knockdown results in upregulation of *daf- 21p::mCherry* transgene.
table S2. *C. elegans* strains used in this study.
table S3. Primers used in qPCR analyses.
table S4. Transgenes tested for derepression with *div-1(RNAi)*.

REFERENCES AND NOTES

1. W. G. Kelly, A. Fire, Chromatin silencing and the maintenance of a functional germline in *Caenorhabditis elegans*. *Development* **125**, 2451–2456 (1998).
2. B. Lehner, A. Calixto, C. Crombie, J. Tischler, A. Fortunato, M. Chalfie, A. G. Fraser, Loss of LIN-35, the *Caenorhabditis elegans* ortholog of the tumor suppressor p105Rb, results in enhanced RNA interference. *Genome Biol.* **7**, R4 (2006).
3. B. D. Towbin, C. González-Aguilera, R. Sack, D. Gaidatzis, V. Kalck, P. Meister, P. Askjaer, S. M. Gasser, Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150**, 934–947 (2012).
4. X. Wu, Z. Shi, M. Cui, M. Han, G. Ruvkun, Repression of germline RNAi pathways in somatic cells by retinoblastoma pathway chromatin complexes. *PLOS Genet.* **8**, e1002542 (2012).
5. L. J. Gaydos, W. Wang, S. Strome, H3K27me and PRC2 transmit a memory of repression across generations and during development. *Science* **345**, 1515–1518 (2014).
6. A. Sapetschnig, P. Sarkies, N. J. Lehrbach, E. A. Miska, Tertiary siRNAs mediate paramutation in *C. elegans*. *PLOS Genetics* **11**, e1005078 (2015).
7. O. Rechavi, L. Houri-Ze'evi, S. Anava, W. S. S. Goh, S. Y. Kerk, G. J. Hannon, O. Hobert, Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell* **158**, 277–287 (2014).
8. E. L. Greer, T. J. Maures, D. Ucar, A. G. Hauswirth, E. Mancini, J. P. Lim, B. A. Benayoun, Y. Shi, A. Brunet, Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. *Nature* **479**, 365–371 (2011).
9. D. Schott, I. Yanai, C. P. Hunter, Natural RNA interference directs a heritable response to the environment. *Sci. Rep.* **4**, 7387 (2014).
10. A. Tauffenberger, J. A. Parker, Heritable transmission of stress resistance by high dietary glucose in *Caenorhabditis elegans*. *PLOS Genet.* **10**, e1004346 (2014).
11. J. K. Arico, D. J. Katz, J. van der Vlag, W. G. Kelly, Epigenetic patterns maintained in early *Caenorhabditis elegans* embryos can be established by gene activity in the parental germ cells. *PLOS Genet.* **7**, e1001391 (2011).
12. A. Klosin, E. Casas, C. Hidalgo-Carcedo, T. Vavouri, B. Lehner, Transgenerational transmission of environmental information in *C. elegans*. *Science* **356**, 320–323 (2017).
13. S. E. Encalada, P. R. Martin, J. B. Phillips, R. Lyczak, D. R. Hamill, K. A. Swan, B. Bowerman, DNA replication defects delay cell division and disrupt cell polarity in early *Caenorhabditis elegans* embryos. *Dev. Biol.* **228**, 225–238 (2000).
14. T. C. Evans, C. P. Hunter, Translational control of maternal RNAs. *WormBook* **10**, 1–11 (2005).
15. J. W. K. Ho, Y. L. Jung, T. Liu, B. H. Alver, S. Lee, K. Ikegami, K.-A. Sohn, A. Minoda, M. Y. Tolstorukov, A. Appert, S. C. J. Parker, T. Gu, A. Kundaje, N. C. Riddle, E. Bishop, T. A. Egelhofer, S. S. Hu, A. A. Alekseyenko, A. Rechtsteiner, D. Asker, J. A. Belsky, S. K. Bowman, Q. B. Chen, R. A.-J. Chen, D. S. Day, Y. Dong, A. C. Dose, X. Duan, C. B. Epstein, S. Ercan, E. A. Feingold, F. Ferrari, J. M. Garrigues, N. Gehlenborg, P. J. Good, P. Haseley, D. He, M. Herrmann, M. M. Hoffman, T. E. Jeffers, P. V. Kharchenko, P. Kolasinska-Zwierz, C. V. Kotwaliwale, N. Kumar, S. A. Langley, E. N. Larschan, I. Latorre, M. W. Libbrecht, X. Lin, R. Park, M. J. Pazin, H. N. Pham, A. Plachetka, B. Qin, Y. B. Schwartz, N. Shoresh, P. Stempor, A. Vielle, C. Wang, C. M. Whittle, H. Xue, R. E. Kingston, J. H. Kim, B. E. Bernstein, A. F. Dernburg, V. Pirrotta, M. I. Kuroda, W. S. Noble, T. D. Tullius, M. Kellis, D. M. MacAlpine, S. Strome, S. C. R. Elgin, X. S. Liu, J. D. Lieb, J. Ahringer, G. H. Karpen, P. J. Park, Comparative analysis of metazoan chromatin organization. *Nature* **512**, 449–452 (2014).
16. L. B. Bender, R. Cao, Y. Zhang, S. Strome, The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in *C. elegans*. *Curr. Biol.* **14**, 1639–1643 (2004).
17. E. C. Andersen, H. R. Horvitz, Two *C. elegans* histone methyltransferases repress *lin-3* EGF transcription to inhibit vulval development. *Development* **134**, 2991–2999 (2007).
18. J. M. Garrigues, S. Sidoli, B. A. Garcia, S. Strome, Defining heterochromatin in *C. elegans* through genome-wide analysis of the heterochromatin protein 1 homolog HPL-2. *Genome Res.* **25**, 76–88 (2015).
19. L. G. Edgar, N. Wolf, W. B. Wood, Early transcription in *Caenorhabditis elegans* embryos. *Development* **120**, 443–451 (1994).
20. A. Klosin, B. Lehner, Mechanisms, timescales and principles of trans-generational epigenetic inheritance in animals. *Curr. Opin. Genet. Dev.* **36**, 41–49 (2016).
21. E. L. Greer, B. Becker, C. Latza, A. Antebi, Y. Shi, Mutation of *C. elegans* demethylase *spr-5* extends transgenerational longevity. *Cell Res.* **26**, 229–238 (2016).
22. K. Siklenka, S. Erkek, M. Godmann, R. Lambrot, S. McGraw, C. Lafleur, T. Cohen, J. Xia, M. Suderman, M. Hallett, J. Trasler, A. H. F. M. Peters, S. Kimmins, Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science* **350**, aab2006 (2015).
23. H. Gaillard, T. García-Muse, A. Aguilera, Replication stress and cancer. *Nat. Rev. Cancer* **15**, 276–289 (2015).
24. S. Ruiz, A. J. Lopez-Contreras, M. Gabut, R. M. Marion, P. Gutierrez-Martinez, S. Bua, O. Ramirez, I. Olalde, S. Rodrigo-Perez, H. Li, T. Marques-Bonet, M. Serrano, M. A. Blasco, N. N. Batada, O. Fernandez-Capetillo, Limiting replication stress during somatic cell reprogramming reduces genomic instability in induced pluripotent stem cells. *Nat. Commun.* **6**, 8036 (2015).
25. P. Sarkies, C. Reams, L. J. Simpson, J. E. Sale, Epigenetic instability due to defective replication of structured DNA. *Mol. Cell* **40**, 703–713 (2010).
26. S. Landais, C. D'Alterio, D. L. Jones, Persistent replicative stress alters polycomb phenotypes and tissue homeostasis in *Drosophila* melanogaster. *Cell Rep.* **7**, 859–870 (2014).
27. T. Iida, H. Araki, Noncompetitive counteractions of DNA polymerase ε and ISW2/yCHRAC for epigenetic inheritance of telomere position effect in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **24**, 217–227 (2004).
28. G. Singh, A. J. S. Klar, Mutations in deoxyribonucleotide biosynthesis pathway cause spreading of silencing across heterochromatic barriers at the mating-type region of the fission yeast. *Yeast* **25**, 117–128 (2008).
29. Z. Jasencakova, A. N. D. Scharf, K. Ask, A. Corpet, A. Imhof, G. Almouzni, A. Groth, Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol. Cell* **37**, 736–743 (2010).
30. M. Brauchle, K. Baumer, P. Gönczy, Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in *C. elegans* embryos. *Curr. Biol.* **13**, 819–827 (2003).
31. A. H. Holway, S.-H. Kim, A. La Volpe, W. M. Michael, Checkpoint silencing during the DNA damage response in *Caenorhabditis elegans* embryos. *J. Cell Biol.* **172**, 999–1008 (2006).
32. S. van den Heuvel, Cell-cycle regulation. *WormBook* **2005**, 1–16 (2005).
33. E. Casas, T. Vavouri, Sperm epigenomics: Challenges and opportunities. *Front. Genet.* **5**, 330 (2014).
34. Z. Jasencakova, A. Groth, Replication stress, a source of epigenetic aberrations in cancer? *Bioessays* **32**, 847–855 (2010).
35. A. Burga, M. O. Casanueva, B. Lehner, Predicting mutation outcome from early stochastic variation in genetic interaction partners. *Nature* **480**, 250–253 (2011).
36. B. Lehner, J. Tischler, A. G. Fraser, RNAi screens in *Caenorhabditis elegans* in a 96-well liquid format and their application to the systematic identification of genetic interactions. *Nat. Protoc.* **1**, 1617–1620 (2006).
37. R. S. Kamath, J. Ahringer, Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* **30**, 313–321 (2003).

38. R. S. Kamath, M. Martinez-Campos, P. Zipperlen, A. G. Fraser, J. Ahringer, Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol.* **2**, RESEARCH0002 (2001).
39. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}_{T}$ method. *Methods* **25**, 402–408 (2001).
40. J. T. Wang, J. Smith, B.-C. Chen, H. Schmidt, D. Rasoloson, A. Paix, B. G. Lambrus, D. Calidas, E. Betzig, G. Seydoux, Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *eLife* **3**, e04591 (2014).
41. D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S. L. Salzberg, TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
42. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
43. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

Acknowledgments: All sequencing experiments were performed in the Centre for Genomic Regulation (CRG) Genomics Core Facility. **Funding:** This work was supported by a European Research Council Consolidator grant (616434), the Spanish Ministry of Economy and Competitiveness (BFU2011-26206 and "Centro de Excelencia Severo Ochoa 2013-2017" SEV-2012-0208), the AXA Research Fund, the Bettencourt-Schueller Foundation, Agencia de Gestio d'Ajuts Universitaris i de Recerca (AGAUR), FP7 project 4DCellFate (277899), and the European Molecular Biology Laboratory–CRG Systems Biology Program. A.K. was partially supported by a la Caixa Fellowship. E.C. and T.V. are supported by the Spanish Ministry of Economy and Competitiveness (grants BFU2015-70581 and IEDI-2016-00787). E.C. is supported by an FI AGAUR PhD fellowship. **Author contributions:** A.K., K.R., and C.H.-C. performed the experiments and data analyses. E.C. and T.V. designed and performed analyses of RNA sequencing data. A.K. and B.L. designed the experiments, interpreted the data, and wrote the manuscript. **Competing interests:** The authors declare they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. RNA sequencing data are available at Gene Expression Omnibus under the GSE95034 entry number.

Submitted 3 April 2017
Accepted 21 July 2017
Published 16 August 2017
10.1126/sciadv.1701143

Citation: A. Klosin, K. Reis, C. Hidalgo-Carcedo, E. Casas, T. Vavouri, B. Lehner, Impaired DNA replication derepresses chromatin and generates a transgenerationally inherited epigenetic memory. *Sci. Adv.* **3**, e1701143 (2017).