Global Hypertranscription in the Mouse Embryonic Germline

Graphical Abstract

Highlights
- Cell-number-normalized methods reveal global hypertranscription in vivo
- mRNAs for biosynthesis, rRNA, and transposon RNAs are amplified in mouse PGCs
- Hypertranscription correlates with cell size and translation
- Myc factors and P-TEFb promote PGC hypertranscription

Authors
Michelle Percharde, Priscilla Wong, Miguel Ramalho-Santos

Correspondence
mrsantos@ucsf.edu

In Brief
Percharde et al. find that mouse primordial germ cells are in a state of hypertranscription driven by Myc factors and P-TEFb.

Accession Numbers
GSE89711
Global Hypertranscription in the Mouse Embryonic Germline

Michelle Percharde, Priscilla Wong, and Miguel Ramalho-Santos

INTRODUCTION

The germline carries developmental totipotency on to the next generation, as well as genetic and epigenetic predispositions to disease. Primordial germ cells (PGCs), the embryonic precursors to mature oocytes and spermatozoa, are thus essential for ontogeny and species survival.

In mice, PGCs arise from a small group of epiblast cells that are first identifiable at embryonic day (E)7.25. From E7.75 to E10.5, these cells migrate to the developing gonads where they proliferate and, from E12.5, commence sexual differentiation (Tam and Snow, 1981; Western et al., 2008). PGC development is accompanied by a remarkable level of epigenetic reprogramming, such as genome-wide DNA demethylation, including the removal of genomic imprints, widespread changes to histone modifications and variants, and X chromosome reactivation in females (reviewed in Saitou et al., 2012). Significant insights have been gained into the unique chromatin reprogramming that PGCs undergo during their development. In parallel, several groups, including our own, have investigated the transcriptional profile of PGCs (e.g., Grskovic et al., 2007; Kurimoto et al., 2008; Seisenberger et al., 2012; Sharov et al., 2003). These studies revealed that PGCs express a pluripotency program with a high similarity to that of embryonic stem cells (ESCs), along with unique modules specific to PGCs related to migration and sexual differentiation.

RESULTS

Increased RNA Polymerase I and II Transcripts in PGCs

To investigate the transcriptome of PGCs and neighboring somatic cells of the gonad (soma), we first took advantage of the Oct4/GFP transgenic mouse line (Yeom et al., 1996). PGCs and somatic cells by their transcriptional output. Our results document a dramatic upregulation of the PGC transcriptome in male and female E13.5 gonads, inclusive of many genes associated with increased biosynthetic capacity and growth. Moreover, we find that PGC hypertranscription depends on the activity of Myc/Max and pTEFb, highlighting a potentially important role for these proteins in germ cell development. These results demonstrate that global hypertranscription occurs in vivo and reveal that the PGC transcriptome is much more dynamic than previously thought.

SUMMARY

Primordial germ cells (PGCs) are vital for inheritance and evolution. Their transcriptional program has been extensively studied and is assumed to be well known. We report here a remarkable global upregulation of the transcriptome of mouse PGCs compared to somatic cells. Using cell-number-normalized genome-wide analyses, we uncover significant transcriptional amplification in PGCs, including mRNAs, rRNA, and transposable elements. Hypertranscription preserves tissue-specific gene expression patterns, correlates with cell size, and can still be detected in E15.5 male germ cells when proliferation has ceased. PGC hypertranscription occurs at the level of nascent transcription, is accompanied by increased translation rates, and is driven by Myc factors n-Myc and l-Myc (but not c-Myc) and by P-TEFb. This study provides a paradigm for transcriptional analyses during development and reveals a major global hyperactivity of the germline transcriptome.
and soma were profiled at E13.5, when sexual differentiation is readily apparent and cells are entering cell-cycle arrest (Bowles and Koopman, 2007). Quantification of total RNA from equal numbers of sorted PGCs or soma revealed approximately 4-fold higher amounts of RNA in male PGCs than in soma and 3-fold in female PGCs (Figure 1A). We next performed cell-number-normalized (CNN) qRT-PCR for common housekeeping genes. Relative to soma, PGCs exhibit significantly higher levels of housekeeping genes typically used for qRT-PCR normalization, including Gapdh, Ubb, Rpol7, and H2A (Figure 1B). Elevated gene expression is apparent throughout PGC development from E9.5 to E13.5 (Figures S1A and S1B), confirming that this phenomenon occurs over a wide window of PGC development.

Moreover, PGC transcription is elevated in comparison to a range of E13.5 tissues and is comparable to cultured ESCs (Figure S1C). Next, to rule out increased mRNA stability in PGCs compared to soma, the expression of both primary (unspliced) and mature RNA polymerase (RNA Pol) I and II transcripts was profiled at E13.5. CNN qRT-PCR confirmed that both pre-rRNA (Figure 1C) and primary mRNAs (Figure 1D) are significantly upregulated in male and female E13.5 PGCs. Interestingly, both male and female PGCs are larger than somatic cells and similar to ESCs (Figures S1D and S1E). Taken together, these results indicate that both RNA Pol I and II transcripts are elevated in male and female PGCs.

**Upregulation of Nascent PGC Transcription and Translation**

The increase in primary RNA transcripts in PGCs (Figures 1C and 1D) suggested that the rate of nascent transcription might be higher in the germline compared to soma. To test this, we performed nascent RNA-labeling experiments in wild-type E13.5 embryos. Cells from dissociated male or female gonads were processed for 5-ethynyl-uridine (EU) incorporation, allowing quantification of nascent transcription by flow cytometry (Koh et al., 2015). We compared the EU incorporation of Ddx4/MVH (mouse Vasa homologue)-positive male or female PGCs with MVH-negative soma or limb (Figure S1F). This revealed a significant increase in nascent transcription in both male and female PGCs (Figures 1E and 1F). Moreover, the rate of nascent transcription in soma is very similar to that in E13.5 limb cells, confirming that hypertranscription is due to a specific elevation of transcriptional activity in PGCs and not a soma-specific suppression (Figure 1E). Finally, we asked whether PGC hypertranscription is linked to elevated protein synthesis. Nascent transcription assays measuring incorporation of L-homopropargylglycine (HPG) revealed that translation is also significantly higher in PGCs, relative to somatic cells (Figures 1G and 1H). Overall, these results reveal a significant elevation in RNA and protein synthesis in PGCs.

**Preservation of Hypertranscription in Male E15.5 GCs**

Examples of hypertranscription in other developmental contexts are often associated with augmented proliferation (Guzman-Ayala et al., 2015; Koh et al., 2015; Nie et al., 2012). We next asked whether germ cells that are not proliferating still exist in a state of hypertranscription. We analyzed E15.5 embryos, when both male and female germ cells (GCs) have entered cell-cycle arrest (Bowles and Koopman, 2007). EU incorporation experiments revealed that total nascent transcription is reduced in male and female GCs at E15.5 compared to E13.5 and that there is more heterogeneity in the transcriptional output of GCs at E15.5 (Figures 2A and 2B). Nevertheless, male E15.5 GCs, overall, display levels of nascent transcription 1.5- to 2-fold higher than soma and limb (Figures 2A and 2B; Figure S2A). In agreement with these results, the levels of RNA Pol II Ser-2P, which marks elongating polymerase, are higher in PGCs of both sexes at E13.5 and remain higher in male GCs at E15.5 (Figures 2C and S2B–S2D). These results indicate that GC hypertranscription persists in male embryos, even after proliferation has ceased.

**Global Analysis of Hypertranscription by CNN RNA-Seq in E13.5 PGCs**

The EU data show that nascent transcription is globally elevated in PGCs but do not resolve which genes are transcriptionally amplified. Are most/all transcripts globally elevated in the developing germline, or do PGCs only express a subset of genes at very high levels? To answer this question, we performed CNN RNA sequencing (RNA-seq) at E13.5 (Tables S1 and S2). Traditional RNA-seq experiments cannot provide accurate differences in expression per cell (Lovén et al., 2012; Percharde et al., 2017), since libraries are created from equal amounts of RNA, and data are normalized to read depth. For CNN analysis, we isolated RNA from equal numbers of male and female PGCs or soma from Oct4/GFP mice and added synthetic External RNA Controls Consortium (ERCC) RNA spike-ins according to cell number prior to library generation (Figure 3A) (ERCC, 2005). Correlation analyses confirmed that PGCs and soma are transcriptionally distinct from each other and are separated by sex, as expected (Figure S3A). We next asked whether the total number of expressed genes varies between PGCs and soma. Interestingly, all cell types express a similar number of genes (Figure 3B). In contrast, the number of highly expressed genes is higher in male and female PGCs than in soma (Figure 3B). In agreement, transcripts in PGCs show an increase in median gene expression, compared to soma (Figure S3B), paralleling the total nascent transcription data (Figure 1E). These results indicate that PGCs express the same approximate number of genes as soma cells, but many genes are expressed at elevated levels.

To determine the expression patterns of individual genes, we generated heatmaps from read-depth normalized or CNN RNA-seq data. In contrast to traditional read-depth normalization, CNN analysis demonstrates that the vast majority of genes are expressed at higher levels in PGCs than soma in both sexes (Figure 3C, Class 1) or in a sex-specific manner (Figure 3C, Class 2 and Class 3). In contrast, much fewer genes are upregulated in soma (Figure 3C, Class 4). Supporting this analysis, 36% and 59% of genes are significantly upregulated more than 4-fold in male and female PGCs, respectively, but only 14% and 8% are upregulated in soma (Figure 3D; Table S1). Hypertranscription preserves tissue specificity in that PGCs do not upregulate somatic developmental regulators, such as Hox genes (Figures 3C and S3C). Importantly, such global amplification of the transcriptome is not apparent without CNN analysis (Figure 3C, left), highlighting the need...
Figure 1. Transcriptional Output of RNA Pol I and Pol II Is Elevated in PGCs
(A) Quantification of total RNA per cell in PGCs or soma. M/F, male/female.
(B and C) CNN qRT-PCR analysis of (B) housekeeping genes or (C) pre-rRNA in PGCs or soma.
(D) Comparison of mature versus primary mRNA transcript expression in PGCs or soma.
(E) Representative histograms of nascent transcription (EU incorporation) in E13.5 gonads.
(F) Quantification of EU data, normalized to limb in each experiment. Fluor., fluorescence.
(G and H) Same as in (E) and (F), respectively, but for nascent translation (HPG incorporation).
All panels are representative of three or more biological replicates; data are indicated as mean ± SEM. *p < 0.05; **p < 0.01; ***p < 0.001, NS, not significant, two-tailed Student’s t test.
See also Figure S1.
for appropriate normalization methods to probe global transcriptomic changes.

We next asked what types of genes are subject to hypertranscription in PGCs. Gene Ontology (GO) analysis of Class 1 genes includes many significant terms for ribosome biogenesis and rRNA processing (Figure S3D). Strikingly, all ribosomal protein genes are upregulated in male and female PGCs (Figure 3E). These results further highlight that the biosynthetic capacity of PGCs is, overall, elevated relative to soma, in agreement with their increased protein synthesis (Figures 1G and 1H), and cellular hypertrophy (Figures S1D and S1E).

Next, we investigated whether CNN analysis reveals distinct patterns of transposable element (TE) expression in PGCs. TEs are thought to be under tight control in the germline so as to preserve genomic integrity. Surprisingly, we found that nearly all TE families are more highly expressed in PGCs than soma.
Hypertranscription has been understudied but is likely pervasive in stem and progenitor cells during development (Percharde et al., 2017). Here, we provide transcriptome-wide documentation of hypertranscription in vivo. We show that the mouse embryonic germline sustains a remarkable level of hypertranscription relative to that of somatic cells at the same stage. This phenomenon can be uncoupled from cell proliferation and is correlated with increased translation and cell size. Importantly, the extent of hypertranscription in the germline only becomes clear after using CNN approaches. This work provides a set of experimental tools to investigate hypertranscription in vivo that will be of broad applicability in other contexts. The results raise several interesting questions as to the regulation and function of hypertranscription in the germline.

Among the most highly upregulated sets of genes in PGCs are those for ribosome biogenesis and translation. This suggests that increased transcription and translation may promote the increase in cell size in PGCs, in agreement with previous accounts of hypertrophy arising due to elevated biosynthesis (Kim et al., 2000). Indeed, an increase in nuclear size has been reported to occur in PGCs from E9.5 to E11.5 (Hajkova et al., 2008; Kagiwada et al., 2013). Moreover, feedback mechanisms whereby a larger cell volume reinforces elevated transcription may exist and contribute to propagating a state of increased biosynthesis (Padovan-Merhar et al., 2015). Interestingly, while female E15.5 GCs appear to have ceased hypertranscription, they are still larger than soma (Figure S2D). This likely reflects a very recent decrease in transcriptional activity, potentially coupled with entry into meiosis, which occurs in female, but not male, GCs at this stage. A loss in hypertranscription may be due to reduced Max expression, which is already lower than in males at E13.5 (Figures 4B and S4A). In agreement, it has been proposed that a reduction in the level of Max is necessary for cultured female PGCs to enter meiosis (Maeda et al., 2013).
Figure 4. Myc Factors and P-TEFb Promote PGC Hypertranscription

(A) Representative genome browser views of RNA-seq data for cMyc, Mycn, and Mycl in E13.5 PGCs and soma.
(B) CNN qRT-PCR validation of the expression of Myc factors in male and female PGCs (MPGC and FPSC, respectively) or soma (Msoma and Fsoma, respectively). Data for each gene are normalized to its expression level in male soma and represented as mean ± SEM for three independent biological replicates.
(C) Schematic of small-molecule modulation of nascent transcription in dissociated gonads.
(D) Representative histograms and quantification of EU incorporation in PGCs and soma incubated with DMSO or 50 μM 10058-F4, an inhibitor of Myc function. (E) Data as in (D), except that 2 μM flavopiridol (FP), an inhibitor of P-TEFb, was used instead of 10058-F4. Data for (D) and (E) are representative of three to four independent biological replicates. *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; ns, not significant, two-tailed Student’s t test. See also Figure S4.
We show that hypertranscription is not limited to unique genes and rDNA but includes TE s: on a per-cell basis, transcripts from all major retrotransposon families are detected at higher levels in PGCs than in soma. This is surprising in light of studies of epigenetic repressive marks in PGCs, which imply that the expression of TE s is heavily suppressed in the germline (e.g., Bourc’his and Bestor, 2004; Liu et al., 2014). Instead, our data indicate that, while under the control of mechanisms that prevent transposition and preserve genomic integrity, TE RNAs are still abundant in PGCs. It will be of interest to investigate whether TE RNAs play novel regulatory roles in the germline (Gerdes et al., 2016).

Another important finding from this study is the role of Myc proteins in promoting PGC hypertranscription. This is likely driven by n-Myc and/or l-Myc, as c-Myc is repressed by Blimp1 in PGCs (Magnúsdóttir et al., 2013). The major mechanism by which Myc-stimulated transcriptional amplification occurs in cancer cells is P-TEFb-regulated transcriptional pause release (Rahl et al., 2010), and our data suggest that the same mechanism is at play in PGCs. PGC-specific deletion of Myc genes, Max, or regulators of transcriptional pause release may shed light on the mechanisms that drive hypertranscription in the mouse germline.

In spite of a widespread amplification of the transcriptome, PGC hypertranscription maintains tissue-specific gene expression patterns, given that PGCs do not express transcriptional regulators of the development of somatic lineages. The silencing of somatic genes in the context of PGC hypertranscription is all the more striking, given the genome-wide loss of DNA methylation (Seisenberger et al., 2012), an epigenetic mark that often represses gene expression. It has been shown that the promoters of somatic developmental regulators are depleted in binding sites for activating transcription factors such as Myc (Ku et al., 2008) and are marked with H3K4me3/H3K27me3 bivalent chromatin in PGCs (Sachs et al., 2013). Thus, a combination of the architecture of the Myc transcriptional network with repressive histone marks such as H3K27me3 likely ensures tissue specificity in the context of PGC hypertranscription.

PGC hypertranscription is likely to be subject to strict temporal control. Immunofluorescence (IF) data from Seki et al. (2007) suggest that RNA Pol II activity is suppressed in E8.5 PGCs but increases thereafter. Our own data indicate that PGC hypertranscription is ongoing between E9.5 and E15.5, peaking around E11.5, and has ceased in females by E15.5. Thus, hypertranscription in PGCs coincides with migration, colonization of the gonadal niche, and early stages of differentiation in the gonads. Our results raise the question of how hypertranscription might regulate germline development. Migratory and early gonadal PGCs undergo rapid growth and proliferation, which may depend on the ability of Myc factors to sustain high levels of hypertranscription and biomass accumulation. Another potential function of hypertranscription is to drive cell competition. The observed heterogeneity in the transcriptional and translational output of PGCs (Figures 1E–1H and 2A) and the underlying role for Myc factors (Figure 4) raise the possibility that hypertranscription may mediate competition between PGCs (Perchard et al., 2017). There is evidence for competition between GCs for the somatic niche in colonial ascidians (Laird et al., 2005), the Drosophila germline (Extaavour and García-Bellido, 2001), and postnatal mouse testis (Shinohara et al., 2002). Moreover, GC competition in the Drosophila ovary has been shown to be regulated by high levels of Myc (Rhiner et al., 2009), which may also drive competition in the post-implantation mouse epiblast (Clavería et al., 2013; Sancho et al., 2013). Selection of PGCs based on their biosynthetic output could be a conserved mechanism to ensure that the cells most fit to support embryonic development go on to contribute to the next generation. Overall, our data warrant future studies of hypertranscription during development of other lineages, as well as dissection of its roles in germline biology.

**EXPERIMENTAL PROCEDURES**

**Animal Work and PGC Isolation**

E13.5 or E15.5 embryos were derived from wild-type 6- to 8-week-old ICR females mated either to ICR or Oct4/GFP B6 males (Yeom et al., 1996). The morning of the day of detection of the copulatory vaginal plug was designated as E0.5. All animal experiments were conducted in accordance with the guidelines of the University of California, San Francisco (UCSF), Institutional Animal Care and Use Committee, protocol AN091331-03.

**CNN qRT-PCR**

E13.5 Oct4/GFP gonads were pooled and enzymatically digested in 0.5% trypsin and 0.8 mg/mL DNase I (Worthington Biochemical). PGCs and matched soma were isolated utilizing a BD FACSAria II (BD Biosciences) according to GFP expression. Equal numbers of cells were sorted directly into Buffer RLT. RNA was purified, and DNase was treated and then used to generate cDNA for qRT-PCR.

**Nascent Transcription and Translation Assays**

Gonads from wild-type ICR embryos were enzymatically dissociated and incubated with 1 mM EU or 50 μM HPG at 37°C for 1 hr. Where indicated, DMSO, 10058-F4 (50 μM, Sigma F3055), or flavopiridol (2 μM, Sigma F3680) was included 30 min prior to the addition of EU, for a total 90-min incubation. Nascent RNA was labeled with Click-IT RNA imaging kits (Thermo Fisher Scientific), followed by incubation with anti-Mvh (Abcam, ab13840) to stain PGCs and GCs, as described previously (Wake et al., 2013).

**CNN RNA-Seq**

10,000 PGCs or 30,000 soma were purified from pooled Oct4/GFP embryonic gonads and used to isolate RNA as described earlier. 4 μL dilution ERCC control mix (ERCC, 2003) (1:10,000, Thermo Fisher Scientific) was added per 10,000 cells to RNA prior to library generation. Libraries were created from 8–10 ng DNase-I-treated total RNA, using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs), pooled at equimolar concentrations, and sequenced on an Illumina HiSeq 4000. Three replicates were sequenced per condition, each yielding >10⁶ single-end 50-bp reads. Detailed bioinformatic methods are available in the Supplemental Experimental Procedures.

**Statistical Analyses**

All statistical calculations were performed with GraphPad 7.0 software; details of individual tests are outlined within each figure legend.

**ACCESSION NUMBERS**

The accession number for the RNA-seq data reported in this paper is GEO: GSE89711.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.05.036.
AUTHOR CONTRIBUTIONS

M.P. and M.R.-S. conceived of the project. M.P. and P.W. performed dissections, flow cytometry, nascent Click-IT assays, and immunofluorescence imaging and analyses. P.W. performed cryosections, immunofluorescence, and flow-cytometry analyses. M.P. designed and performed all other bench experiments and performed the bioinformatics analyses. M.P. and M.R.-S. wrote the manuscript, with input from P.W.

ACKNOWLEDGMENTS

We are grateful to Diana Laird, Marco Conti, Robert Bieloch, Jacob Freimer, and members of the Santos lab for input or critical reading of the manuscript. We thank Robert Bieloch and Amy Chen for mir290mCherry/mir302GFP ESCs and Steffen Biechele for trophoblast stem (TS) cells. This work was supported by CIRM postdoctoral fellowship TG2-01153 to M.P. and by NIH grants R01OD012204 and R01GM113014 to M.R.-S.

REFERENCES


