Fundamental differences in endoreplication in mammals and Drosophila revealed by analysis of endocycling and endomitotic cells

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Fundamental differences in endoreplication in mammals and Drosophila revealed by analysis of endocycling and endomitotic cells

Noa Shep, Jessica R. Von Stetina, George W. Bell, Shinobu Matsura, Katya Ravid and Terry L. Orr-Weaver

Throughout the plant and animal kingdoms, cells in specific tissues increase their DNA content to become polyploid, making an understanding of the mechanisms involved in polyploidization a fundamental question in biology. The biological rationale for increasing genome content during differentiation appears most commonly to be to increase cell size, although it may facilitate increased cell metabolism. Much remains to be discovered in terms of how genomic DNA content is increased during differentiation and how this affects gene expression.

The cell cycle alterations leading to polyploidy have been investigated in Drosophila and mammals, revealing diverse mechanisms. The endocycle is the modified cell cycle most variant from the canonical division cycle, consisting solely of alternating synthesis (S) and gap (G) phases. In Drosophila, nearly all differentiated tissues increase DNA content via the endocycle (3). The ploidy of differentiated Drosophila cells range from 16C (where C is the haploid genome content) to 1024C, depending on the cell type (4, 5). The structure of the chromosomes in most Drosophila cells is polypotent, differing from polyploid cells in that all replicated copies of the sister chromatids are physically aligned, giving a consistent banding pattern. Some cell types, such as mammalian megakaryocytes (MKs), undergo endomitosis, a modified cell cycle with G1, S, and G2 phases but only a partial M phase in that nuclear division and cytokinesis do not occur (for review see ref. 6). Thus, the cells produced by endomitosis are mononucleated, like those from the endocyte, but the replicated sister chromatids are separated and polyploid rather than polytene. A third mechanism leading to polyploidy involves nuclear division but no cytokinesis to produce multinucleate cells, as in tetraploid hepatocytes (7).

In Drosophila endocycles, differential DNA replication occurs, rather than an integral doubling of the genome during S phase (8). This can lead to increased gene copy number, gene amplification, of specific genomic regions, as occurs at six sites in ovariian follicle cells (9). More commonly, underreplication leads to reduced gene copy number. All Drosophila polypotent cells examined to date have a highly reduced copy number of heterochromatin as well as underreplicated euchromatic regions. These underreplicated regions outside of heterochromatin blocks can be either tissue specific or common to several tissues (10). Whereas gene amplification can be a strategy for robust expression of specific genes over a short developmental time, underreplication may conserve resources by avoiding replication of gene-poor regions of the genome.

Mammalian MKs achieve a ploidy of 128C, and the resulting increased cell size is necessary for sufficient platelet production (6). During MK endomitosis, anaphase A chromatid separation occurs, but the spindle does not elongate, and mitosis is aborted without nuclear division (11). Lagging chromosomes are observed during anaphase, possibly reflecting incomplete DNA replication, although this has not been investigated.

The trophoblast giant cells (TGCs) provide a barrier between the maternal blood supply and the embryo proper, facilitated by the large size of the cells (12–14). In rodents, they become highly polyploid; reaching up to 512C in mouse (15). The murine TGCs differ from syncytiotrophoblasts found in both rodents and human in that they contain a single, highly polyploid nucleus, whereas the latter arise by cell fusion of G1 cells, yielding multinucleate cells with diploid nuclei (16, 17). In addition, TGCs are on an endocycle, and polypotent structures have been visualized by in situ hybridization, although in contrast to Drosophila the polypotent structures do not extend the length of each chromosome (18). The TGCs' chromosome morphology is dynamic during development, with the chromosomes appearing to fragment at day 10 of mouse embryogenesis (16). Mouse TGCs can be differentiated in culture from trophoblast stem cells by the absence of fibroblast growth factor 4, and a rat trophoblast cell culture line exists that becomes polyploid in culture, attaining up to 64C.
ploidy (19, 20). Analysis of these models revealed that the cyclin-dependent kinase (CDK) inhibitor p57 is necessary for the onset of the endocycle (21, 22). Differential DNA replication, however, was not examined in these cell culture models.

Here, we investigate the nature of S phase in the modified cell cycles producing MKs and TGCs, using primary MKs differentiated from precursor cells and mouse TGCs isolated directly from implantation sites. Through array-based comparative genome hybridization (aCGH), we systematically analyzed the genome for regions subject to differential DNA replication, uncovering fundamental differences in endoreplication in mouse and *Drosophila*.

Results

**Primary Polyploid Murine Megakaryocytes Undergo Uniform Genome Replication.** Differential replication is a phenomenon common to all *Drosophila* polyploids and polyploid tissues (8). Before this study, however, it was not known whether mammalian cells that become polyploid also change gene copy number in parts of the genome. aCGH permitted quantification of gene copy number across the euchromatic portion of the genome. We first investigated mouse bone marrow MKs, which polyploidize via endomitosis (Fig. 1A), the modified cell cycle least variant from the archetypal cell cycle. Recent studies suggested that low-ploidy (2C, 4C) murine MKs have uniform replication of the genome, given the relatively normal anaphase figures (23). In contrast, higher ploidy MKs could have differential replication that may contribute to the observed aberrant mitotic figures (23, 24).

We obtained polyploid MKs for aCGH by isolating bone marrow progenitor cells from C57BL/6j mice and culturing them in the presence of thrombopoietin (TPO) for 4 d to induce MK polyploidization. Following BSA gradient fractionation, we obtained populations of 2C and 16C–128C MKs (Fig. 1B). To validate the purity of our polyploid-enriched MK fraction, we analyzed DNA content by fluorescence-activated cell sorting (FACS) of total MKs (~256C (34, 35), which is in agreement with our day 9.5 (E9.5) implantation sites from C57BL/J6 mice (32), as cells at this stage are still in a layer amenable to microdissection. Furthermore, we manually removed the portion of the TGCs in direct contact with the spongiosotrophoblast layer and the labyrinth layer to avoid collecting any polyploid cells from the former or multinucleated syncytiotrophoblast cells from the latter (12). We then analyzed the ploidy of the isolated TGC population, measuring DAPI intensity as described (33) and normalizing to diploid embryonic DNA content. TGCs had ~80-fold (median = 82) more DNA than embryonic cells (Fig. 2A), indicating that the majority of the cells had undergone roughly six (log_2 128C) endocycles. It has been estimated that the maximal ploidy level at day 9 is 128–256C (34, 35), which is in agreement with our day 9.5 DAPI quantifications. RNA-Seq quantification (RNA-Seq) identification of the transcriptome from these isolated TGCs further supported the purity of the cells (see below and Fig. S4).

**Trophoblast Giant Cells Do Not Undergo Significant Differential DNA Replication.** Genomic DNA was isolated from the purified TGCs to perform aCGH to search for over- or underreplicated genomic intervals are resistant to replication and are underreplicated. In addition, genomic regions harboring some of the genes uniquely or selectively expressed abundantly in MKs, such as glycoprotein 6 (GPVI/Gp6) (25), c-myeloproliferative leukemia virus oncogene (c-Mpl) (26, 27), cyclin D3 (28), and cyclin E (29), do not have amplified copy number above the overall ploidy of the cells (Fig. S2).

**Microdissection of a Pure Population of Embryonic Day 9.5 Murine Trophoblast Giant Cells.** TGCs are similar to *Drosophila* polyploid cells in that they use the endocycle to become polyploid and thus possibly undergo gene amplification or underreplication (16). Parietal TGCs (P-TGCs, termed TGCs for simplicity from here on) (30, 31) form a thin layer of extraembryonic cells separating the maternally derived decidua from the embryonically derived layers of the murine placenta (Fig. 2A), with placental morphology and TGC accessibility drastically changing throughout gestation (12, 13). We chose to isolate TGCs rather than analyze cell culture models to avoid potential differences in the endocycle that may occur in cell culture and also because the native TGCs attain a higher level of ploidy.

To obtain DNA and RNA with minimal contamination from the surrounding layers, we microdissected TGCs from embryonic day 9.5 (E9.5) implantation sites from C57BL/J6 mice (32), as cells at this stage are still in a layer amenable to microdissection. Further, we manually removed the portion of the TGCs in direct contact with the spongiosotrophoblast layer and the labyrinth layer to avoid collecting any polyploid cells from the former or multinucleated syncytiotrophoblast cells from the latter (12). We then analyzed the ploidy of the isolated TGC population, measuring DAPI intensity as described (33) and normalizing to diploid embryonic DNA content. TGCs had ~80-fold (median = 82) more DNA than embryonic cells (Fig. 2B), indicating that the majority of the cells had undergone roughly six (log_2 128C) endocycles. It has been estimated that the maximal ploidy level at day 9 is 128–256C (34, 35), which is in agreement with our day 9.5 DAPI quantifications. RNA-Seq quantification (RNA-Seq) identification of the transcriptome from these isolated TGCs further supported the purity of the cells (see below and Fig. S4).
genomic regions. DNA copy number measurements of E9.5 TGCs relative to E9.5 embryos from two biological replicate aCGH experiments showed no significant (>1.3-fold relative to overall ploidy) differential replication in TGCs (Fig. 2C and Fig. S3), indicating that no euchromatic regions were over- or underrelicated for even one round of endoreplication at this stage.

**Heterochromatin Is Fully Replicated in Polyploid TGCs and MKs.** In *Drosophila*, underreplication is prevalent in both euchromatic (10) and heterochromatic regions (36, 37). The blocks of heterochromatin surrounding the centromeres and accounting for 20–30% of the length of each chromosome arm are not visible in *Drosophila* polytene chromosomes. Cytological studies on the chromosomes of TGCs have yielded conflicting conclusions about whether and to what extent heterochromatin is underreplicated (15, 18). The bridges observed in anaphase in MKs could be due to underreplication of the heterochromatin. Because the microarrays do not contain probes for heterochromatin sequences, we measured copy number changes in heterochromatin in isolated TGCs and MKs by quantitative real-time PCR (qPCR). Using primer sequences derived from bacterial artificial chromosome (BAC) clones, we performed qPCR on three different heterochromatic regions relative to two single-copy genes (actin and tubulin) in two biological replicates. Strikingly, our results demonstrate that both TGCs and MKs undergo full replication of heterochromatin.

**Transcriptome of Trophoblast Giant Cells.** To identify gene expression signatures associated with TGC differentiation and endoreplication, we determined the transcriptome and micro-RNA (miRNA) expression profile from the purified trophoblasts. The transcriptome analysis confirmed a high level of purity for our hand-dissected samples, with high expression of known TGC markers, and negligible expression of genes known to be expressed in nearby tissues (Fig. S4). As expected, among the most abundant protein-coding RNA molecules of the TGCs (top 30 expressed protein-coding genes shown in Table S1) were many hormone genes (such as the prolactin family) and genes involved in steroid biosynthesis [cytochrome P450, family 11, subfamily A, polypeptide 1 (Cyp11a1), steriodogenic acute regulatory protein (STAR), and hydroxy-delta-5-steroid dehydrogenase 3 beta-and steroid delta-isomerase 6 (Hsd3b6)]. The heart and neural crest derivatives expressed transcript 1 (Hand1) gene, an important regulator of TGC differentiation, also was highly expressed (19). Other highly expressed genes included scavenger receptor class B, member 1 (Scarb 1), required for the phagocytic activity of TGCs (38, 39), and the Cdk1 inhibitor p57 (Cdkn1c), required for TGC endoreplication (22).

We analyzed the E9.5 TGC mRNA-Seq profile with the GOriola gene ontology (GO) tool (http://www.ncbi.nlm.nih.gov/pubmed/19192299), which uses an algorithm that predicts the choice of an arbitrary P value cutoff. Consistent with TGCs being highly metabolic steroidogenic cells, strongly expressed genes are enriched in GO terms for intracellular transport, Golgi vesicle transport, and ATP and steroid biosynthesis (Table S2). Among the enriched GO process categories (Table S2) were metabolism, translation, intracellular transport, Golgi vesicle transport, negative regulation of cell death, ATP biosynthetic process, steroid biosynthetic process, angiogenesis, and cell differentiation involved in embryonic placenta development. The enriched expression of genes involved in ATP biosynthesis and the robust expression of mitochondrial genes are consistent with a recently proposed role of mitochondria in polyplodyization of maternal decidual cells (40). GO terms unique to the mitotic cell cycle and apoptosis were less common, in agreement with recent studies showing polyplody cells to have antiapoptotic properties, important in avoiding cell cycle checkpoints (22, 41).

**Expression Levels of Replication Factors Vary Among Different Polyploid Cell Types.** The *Drosophila* larval polytene salivary gland (SG), fat body, and midgut tissues undergo a dramatic

---

**Fig. 2.** Trophoblast giant cells are devoid of differential replication. (A) Illustration of day 9.5 mouse implantation site. In blue is shown the trophoblast giant cell layer that was microdissected for this study. (B) Scatterplot showing ploidy C values from DAPI microdensitometry of day E9.5 trophoblast giant cells (TGCs) and E9.5 embryonic cells (n = 98 for TGCs and n = 68 for embryos). Polytenic TGCs contain 80-fold more DNA content than embryonic diploid cells; the average C value for TGCs is 162C compared with the average embryonic value set at 2C. (C) Two biological replicates of aCGH arrays on chromosome 1 of the mouse genome, comparing DNA levels from the TGCs at day E9.5 relative to diploid embryonic DNA levels.

**Fig. 3.** Heterochromatin in TGCs and MKs is fully replicated. qPCR measurements of DNA copy number at three independent heterochromatic regions relative to two single-copy genes (actin and tubulin). (A) Fold changes in heterochromatin copy number were similar (ratio of 1) between polytenic TGCs and diploid embryonic cells, and between polyplody and diploid MKs (B), indicating no underreplication in these regions. BAC sequence primers used for this analysis covered both centromeric (P-0416-184N24 and P-RP24-110P17) and telomeric (P-CH25-423E5) regions. Error bars, SEM of two independent biological replicates.
reduction in the expression of genes involved in DNA replication (10, 42). It has been proposed that reduced expression of genes required for DNA replication leads to a slowed S phase and failure to replicate genomic regions normally replicated late in S phase (43), possibly leading to underreplication in these tissues.

We analyzed the expression status of S-phase genes in mammalian TGCs and MKs. S-phase gene transcript levels were quantile normalized (Materials and Methods) (Dataset S1), and compared between polytene TGC or polyploid MK cells and diploid control cells. In these analyses we used our RNA-Seq data from TGCs compared with published RNA-Seq data from mouse embryonic brains and published microarray data sets for polyploid versus diploid human MKs (Fig. 4). We compared expression levels of orthologous genes in the DNA replication pathway and cell cycle genes controlling the G1–S transition and S phase. Strikingly, we found that components of the DNA replication machinery including the origin recognition complex (ORC), minichromosome maintenance (MCM), and proliferating cell nuclear antigen (PCNA) genes were strongly expressed in the TGC lineage (Fig. 4). In contrast to TGCs, polyploid MKs exhibited a reduction in the expression of these genes, but the normalized expression levels were higher than in the Drosophila salivary gland (Fig. 4).

In Drosophila, the onset of endoreplication strongly correlates with the transcriptional and translational down-regulation of key M-phase promoting genes (1). To investigate the expression patterns of these mitotic genes in mammalian polytene and polyploid cells, we compared TGCs and MKs (Fig. S5 and Dataset S1). Similar to Drosophila polytene cells, TGCs show a reduction in the expression of mitotic genes compared with diploid control embryonic cells (Fig. S5), which tightly correlates with the nature of the endocycle and is consistent with previously described findings by MacAuley et al. and Hattori et al. (20). Consistently, their ability to enter mitosis, polyploid MKs showed a different expression pattern, with increased expression of M-phase factors relative to TGCs or Drosophila SGs (Fig. S5).

miRNA Profile of Trophoblast Giant Cells. One of the most abundant noncoding RNA transcripts found in our TGC RNA-Seq data set was H19 (Fragments Per Kilobase of transcript per Million reads mapped (FPKM) = 24,368), a tumor-suppressor noncoding RNA gene (44–47). As H19 is not only the first imprinted noncoding transcript discovered, but also is the known precursor for miR-675 (48, 49), we analyzed the miRNA composition of the TGCs by deep sequencing. Although we did see strong expression of miR-675 (Table S3), it was far from the most abundant miRNA. Therefore, it is likely that the high expression of H19 in TGCs functions to maintain high levels of the long noncoding RNA for its tumor-suppressor properties rather than primarily to give rise to miR-675. Importantly, H19 is expressed in human placental extravillous cytotoxophoblast cells (50). These are polyploid cells that invade the uterus in humans and share the expression of many transcription factors, proteases, and cell adhesion molecules with TGCs (30). It is attractive to speculate that H19 plays a role in controlling the extent of invasiveness of the TGC cell population.

The two most highly expressed miRNAs in TGCs were miR-322 and miR-503 (Table S3), recently shown to down-regulate cell division cycle 25A (cdc25A) (51), whose expression in TGCs is indeed extremely low according to our RNA-Seq data (FPKM = 20). Many of the set-7 family members, implicated in promoting terminal differentiation and cell cycle arrest, and in biological functions such as activation of gene expression in euchromatin (52–54), are highly expressed in TGCs. miR-182 is an antiapoptotic miRNA that is abundantly expressed in TGCs, was recently shown to decrease in human placentae of patients with preeclampsia, a disease associated with extravillous cytotoxophoblast cells that are analogous in many respects to the TGCs of the rodent placenta (55). Little is known about the function of other highly expressed miRNAs (e.g., miR-51), making them interesting candidates for future studies. This catalog of miRNAs and miRNAs expressed in E9.5 TGCs will be a resource for further study of trophoblast differentiation and function.

Discussion

In this study, we compared mechanisms producing polyploid cells by analyzing gene expression and genome copy number in two polyploid mammalian cell types, the TGCs and MKs. By directly isolating TGCs from mouse placenta, we were able to obtain cells of high ploidy in their native state to analyze the parameters of DNA replication by quantification of genome copy number. The advantage of this approach over the use of cell culture models was the ability to analyze higher levels of ploidy and to avoid potential S phase alterations arising in cell culture lines. By examining primary MKs differentiated for only 4 d from bone marrow progenitor cells, possible genome aberrations from cell culture lines were eliminated. The transcriptome of the endocycling TGCs described here, compared with published data on endomitotic MKs, supports the notion that repression of the mitotic machinery is an integral component of endocycling cells both in Drosophila and mammals, in contrast to its presence in cells undergoing endomitosis.

Our aCGH studies and qPCR copy number measurements revealed integral genome doublings in mouse TGCs and MKs, and an absence of differential (>1.5-fold) DNA replication. Thus, in these cells, robust gene expression is not accompanied by gene amplification. It is notable that several extensive domains in TGCs appear reproducibly reduced in copy number and are called peaks by the Ringo and MA2C programs if thresholds are set to less than 1.3-fold. The slight copy number reduction may be a consequence of these domains replicating slowly, thus requiring a longer period of S phase to be duplicated. Therefore, in a high percentage of S-phase cells these regions may not yet be replicated.

The genome-wide aCGH analysis provided a quantitative and direct examination of copy number. Previous cytological studies in MKs raised the possibility of underreplication, given the chromosome bridges observed during aberrant anaphases (11). Because TGCs use an endocycle and all examined endocycling cells in Drosophila exhibit underreplication, we hypothesized that at the very least TGCs heterochromatin would be underreplicated. It was
not possible from previous cytological studies to ascertain gene copy number with certainty or to survey the entire euchromatin genome, and results yielded inconsistent conclusions about differential replication (56–58). TGC chromosomes showed detectable heterochromatin staining and the presence of Barr bodies, but there was no increase in the number and size of heterochromatin blocks with increasing ploidy (15). Attempts to quantify DNA content by cytophotometry were consistent with integral genome doublings in some cells, but showed others with DNA values lower than multiples of 2C (15). A restriction landmark genomic scanning study of CpG islands in late gestation rat placental junctional zone tissue, which contains TGCs, indicated that nearly all of the ~1,000 loci tested showed no differential replication (59). In situ hybridization signals with satellite DNA probes indicated replication of the heterochromatin in TGCs (18).

Given the evolutionary conservation of the endocycle and its universal use in the plant and animal kingdoms, it was surprising to observe the absence of differential replication in TGCs compared with its presence in all examined Drosophila tissues. Analysis of the transcriptome of the purified TGCs revealed a striking elevation in the normalized level of expression of genes required for S phase relative to the transcriptome of the Drosophila salivary gland. S-phase genes are also more highly expressed in MKs compared with Drosophila endocycling cells. These observations suggest that there are differences in the parameters of the endocycle between different organisms.

The differences in expression of replication genes support the model (42, 43) that limiting expression of proteins needed for initiation of DNA replication leads to a slower S phase and prevents the duplication of genome regions that are late replicating. This provides an advantage of not replicating gene-poor regions such as heterochromatin, but at the cost of slow endocycles. The endocycles indeed proceed much faster in the mouse TGCs than in Drosophila. The 256C ploidy at day 9.5 is achieved in 4–5 d of embryogenesis (15), in contrast to the Drosophila salivary gland that takes 7 d to achieve a ploidy of 1024C for a genome only 5% the size of the mouse genome (4). We speculate that the developmental window during which TGCs must attain the proper ploidy is short enough that a faster endocycle is required, necessitating high expression of S-phase genes and forcing replication of heterochromatic regions.

A second potential explanation for why TGCs and MKs integrally replicate their genomes in contrast to Drosophila tissues may lie in differences in genomic organization of the euchromatin and heterochromatin in the two species. In Drosophila, nearly all of the heterochromatin is present in large blocks surrounding each centromere, accounting for 20%–30% of the length of each chromosome arm (60). In contrast, in mouse, although there are satellite DNA present in each centromere and telomere, heterochromatin is dispersed throughout the genome (61–63). We have shown that underreplication in Drosophila results both from an absence of origins across a region and from the presence of flanking chromatin that blocks progression of replication forks that initiate outside of an underreplicated region (64). The presence of heterochromatin in large domains may facilitate its underreplication, as such a domain necessitates only one region per side to impede fork progression. Interspersed murine heterochromatin would require many such blocking regions per chromosome.

It should be noted that the maximal ploidy levels of murine TGCs has been estimated to be 512C–850C at embryonic day 11 (16, 34). Thus, the TGCs undergo one or two more rounds of endoreduplication beyond the stage at which our aCGH experiments were performed. It is conceivable that underreplication or gene amplification do not begin until these last one or two S phases. It is not possible to test this experimentally for TGCs because after embryonic day 9.5 the cells are embedded and cannot be isolated by microdissection. Underreplication for only the last two endocycles, however, would at most lead to a fourfold decrease in copy number. In the Drosophila larval tissues, underreplication begins at the first endocycle in embryogenesis (3, 64), although the Drosophila ovarian nurse cells undergo five endocycles with full replication followed by five with underreplication (65). Consequently, we cannot exclude the possibility that a later developmental change in the endocycle in the mouse TGCs could produce differential DNA replication. In addition, recent partial TGC subtype was analyzed; it is still possible that other TGC subtypes (30) do exhibit differential replication.

Increased DNA content is used as a developmental strategy throughout the plant and animal kingdoms. Despite this evolutionary conservation, we found striking differences between Drosophila and mouse in the use of differential DNA replication within the context of endoreduplication. These differences suggest distinct parameters of the endocycle and possible effects of the genomic organization of heterochromatin and euchromatin on underreplication. It will be informative to examine endocycling cells in plants and other animal tissues to determine the association between differential DNA replication and heterochromatin organization and the link with levels of expression of S-phase genes.

Materials and Methods

Isolation of Mouse MKs and FACS. Isolation of mouse bone marrow cells, culture, purification of polyploid MKs, and FACS are detailed in SI Materials and Methods. These procedures were approved by the Institutional Animal Care and Use Committee at Boston University Medical Center.

Isolation of TGC Cells. Day E9.5 trophoblast giant cells were microdissected from C57Bl/6J mice as previously described (32) from two biological replicates. TGCs were isolated from 29 implantation sites from three pregnant mice (replicate one), and from 14 implantation sites from one pregnant mouse (replicate two). Diploid embryonic controls were E9.5 embryos from the dissected implantation sites. These procedures were approved by the Committee on Animal Care at the Massachusetts Institute of Technology.

aCGH. DNA was prepared from diploid (2C) and polyploid (>16C) MK cell fractions or embryonic cells and purified TGCs, labeled, and hybridized to microarrays as described in SI Materials and Methods.

mRNA-Seq. Day E9.5 TGCs were microdissected from nine implantation sites from one pregnant mouse, immediately transferred into TRizol (Invitrogen), and RNA extracted according to the manufacturer’s protocol. Ten micrograms of RNA were processed with an Illumina mRNA Sample Preparation kit. Sequencing was done as described in SI Materials and Methods.

miRNA-Seq. TGCs miRNA sequencing samples were processed using an Illumina miRNA kit as in ref. 66. After adapter trimming, Illumina reads were aligned to mature miRNAs (miRBase Release 16) using Novalign, and reads per miRNA were calculated.

Comparison of Expression Profiles. Expression data (microarrays and short reads) were downloaded, analyzed separately, and replicates were mean summarized. RNA-Seq profiles were generated using mouse TGCs (this study; 36-nl Illumina reads; n = 1), mouse E14.5 fetal heads (67) (National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) GEO series GSE33017; 36-nl Illumina reads; n = 3), Drosophila salivary glands (64) (NCBI GEO GSE33017; 36-nl illumina reads; n = 2), and Kc167 cells (modENCODE 68; modENCODE_2593; 37-nt illumina reads; n = 2) datasets. Microarray expression profiles were generated using published Drosophila salivary gland and brain-disk complexes (42) (NCBI GEO GSE19029; Genepix two-channel microarrays; n = 1) and human 2C and 16C MKs (69) (EBI ArrayExpress E-TABM-133; Agilent two-channel microarrays; n = 2). Mouse–fly ortholog pairs were obtained from Ensembl v67 and linked to RNA-Seq and expression microarray identifiers to obtain gene sets for further analysis.

Real-Time qPCR. Real-time qPCR was performed in two biological replicates with primers against BAC sequences for centromeric (BAC-RP24-184N24 (Y) and BAC-RP24-110P17 (Y)) and telomeric (BAC-CHZS2-42ES (X)) heterochromatic regions in E9.5 TGCs (polyploidy sample) and E9.5 embryos (diploid sample), using Perfecta SYBR Green FastMix (Quanta BioSciences) with an Applied Biosystems 7300 Real-Time PCR machine according to the manufacturer’s instructions. Primers are listed in SI Materials and Methods. PCR values were normalized to levels of two single-copy genes (β-actin and γ-tubulin).

ACKNOWLEDGMENTS. We thank Olivia Risland for assistance with miRNA-Seq; Inma Barrasa and Sumeet Gupta for bioinformatics advice; Tom
Supporting Information

Sher et al. 10.1073/pnas.1304889110

SI Materials and Methods
Primary Culture of Mouse Megakaryocytes, Ploidy Fractionation and FACS. Bone marrow cells were flushed from femurs and tibias of twenty-four 6- to 8-wk-old C57B6/J mice with CATCH buffer (1 mM adenosine, 2 mM theophylline, 0.38% sodium citrate, 5% (vol/vol) FBs in 1× HBSS). Cells were centrifuged for 5 min at 1,200 × g at 4 °C, washed with PBS (without Ca2+ or Mg2+), and filtered through a 70-μm cell strainer. Cells (5–8 million cells per milliliter) were cultured in six-well plates with Iscove’s Modified Dulbecco’s Medium (IMDM) media (Invitrogen; 21056) containing 10% (vol/vol) FBS, 1% penicillin–streptomycin, 4 mM l-glutamine, and 25 ng/mL human thrombopoietin (generous gift of Kirin Pharma Co.) for 4 d at 37 °C, at which point the cells reach a ploidy between 64C and 128C (Fig. 1B). Attempts to culture for more than 4 d resulted in high levels of cell death. Cells were then collected and layered on a discontinuous BSA density gradient [0%/1.5%/3.0% (wt/vol) BSA in PBS] repeated twice, as described (1). Diploid (2C, 1.5% pellet, nearly 100% purity) and polyploid (16C–128C, 3% (wt/vol) BSA pellet, about 90% purity based on cell mass) (1) (see also Fig. 1B) MKs were collected, washed, and used for DNA extraction.

To analyze DNA content by FACS, cells were incubated for 4 d with thrombopoietin (TPO), and total MKs or polyploid MKs from 3% (wt/vol) BSA pellet were resuspended in MACS buffer (Ca2+ and Mg2+-free PBS, 0.5% fatty-acid–free BSA, 2 mM EDTA). Cells were incubated with rat FITC-conjugated anti-CD41 (BD Biosciences; 553848) or rat IgG1k control (BD Biosciences; 554684) for 30 min at 4 °C, and fixed in ethanol for 30 min at 4 °C. Fifteen minutes before FACS analysis, cells were treated with 100 μg/mL RNase A and stained with 0.05 mg/mL propidium iodide (PI) to stain DNA. Sorting was performed using a FACSCalibur flow cytometer and FL1 (530 ± 30 nm; PI) lasers. FACS plots were generated using the BD CellQuest Pro v5.2 software.

aCGH Analysis. Diploid (2C) and polyploid (>8C) MK cell fractions or embryonic cells and purified trophoblast giant cells (TGCs) were lysed in CHIP buffer (50 mM HEPES, pH 7.5; 140 mM NaCl; 1 mM EDTA; 1% Triton X-100; 0.1% Na deoxycholate), and sonicated (Branson Sonifier 250) on ice with four 20-s pulses at a power of 2, duty constant. Cell suspensions were treated with proteinase K (0.2 mg/L) and RNase (0.1 mg/mL) and DNA was isolated following standard phenol/chloroform extraction. Genomic DNA samples were differentially labeled with Invitrogen’s BioPrime Total for Agilent aCGH labeling kit and cohybridized to a 1-million probe custom tiling array designed for C57B5/J (Agilent-027414). Slides were washed and scanned using an Agilent microarray scanner. Using the Ringo package in R (2), probes were median normalized and smoothed by running medians before calling underreplicated or amplified regions, defined by at least five adjacent probes with a fold change (copy number vs. overall ploidy) greater than 1.3. Micro-arrays were also converted into NimbleGen format and processed by MA2C software (3). All aCGH genome browser figures show the log2-transformed probe intensities of polyploid cells relative to diploid controls.

qPCR Primers. Primers were used as follows: BAC-RP24-184N24 (Y), 5′-GCT CTG GCC GCT CCT AGA A-3′ and 5′-TGG GTT CAA GGA ATG GTC AAA-3′; BAC-RP24-110P17 (Y), 5′-CGA ATC CAA GAA CAC ATT AAA GCA-3′ and 5′-TCC CTG CAT CCC TGG AAT AA-3′; BAC-CH254-232E5 (X), 5′-CTC CTG TGT GTG GGT CAG TG-3′ and 5′-ACA CAG CAG ATG TGA GG-3′.

mRNA Seq. High-quality samples were confirmed by the appearance of a smooth smear of fragments with a peak distribution between 150 and 300 bp on an Agilent Bioanalyzer. After subsequent qPCR library quantification, 2–7 μM of linker-ligated DNA was applied to a flow cell using the Illumina Cluster Station fluids device. Thirty-six-base sequencing was performed on an Illumina Genome Analyzer II (GAII) sequencer, according to Illumina’s standard protocols. Images acquired were processed by the bundled Illumina image extraction pipeline v1.6 and aligned to Mus musculus National Center for Biotechnology Information build mm9 using TopHat (4). Gene-level quantification was performed by Cufflinks (5) using Ensembl gene models.

Comparison of Expression Profiles. For RNA-Seq experiments, TopHat v1 was used to map mouse and Drosophila reads to mouse mm9 canonical chromosomes or to fly dm3 canonical chromosomes, respectively. Genes were quantified with Cufflinks (Footprints Per Kilobase of transcript per Million reads mapped (FPKM) = +1, using Ensembl annotations). For expression microarray experiments, probes were median summarized to get gene-level values and produce channel-level intensities. Microarray probe intensities and FPKM = +1 values were log2-transformed and for each desired comparison, gene levels were quantile normalized, and log2 ratios were scaled to median = 0. S-phase and mitotic genes were selected by hand curation.

Fig. S1. (Continued)
Fig. S1. (Continued)
Fig. S1. Array-based comparative genome hybridizations (aCGHs) for all chromosomes of the mouse genome, comparing euchromatic gene copy number between murine polyploid (>8C) and diploid (2C) MKs.
Fig. S2. Close-up of aCGH arrays in regions of murine chromosomes harboring genes that are exclusively expressed or highly abundant in MKs including glycoprotein VI (Gp6), myeloproliferative leukemia virus oncogene (Mpl), cyclin D3 (Ccnd3), and cyclin E1 (Ccne1). No amplification is detected in these regions.
Fig. S3. (Continued)
Fig. S3. (Continued)
Fig. S3. (Continued)
Fig. S3. Two biological replicate aCGH arrays for all chromosomes of the mouse genome, comparing DNA levels from the trophoblast giant cell layer at embryonic day 9.5 (E9.5) with diploid embryonic DNA.
Fig. S4. Hand-dissected TGC samples are a nearly pure population. Shown are wiggle plots from Illumina RNA Sequencing illustrating a high number of reads of genes known to be markers uniquely expressed by TGCs \[\text{prolactin family 3, subfamily d, members 1–3 (Prl3d1–3)},\] whereas genes known to be expressed in the spatially adjacent spongiotrophoblast [\text{distal-less homeobox 3 (DLX3)} (shown), trophoblast specific protein alpha (Tpbpa), Nodal], deciduas and/or ectoplacental cone [\text{prolactin family 6, subfamily a, member 1 (Prl6a1)} (shown), prolactin family 8, subfamily a, member 2 (Prl8a2)] or in the labyrinth [\text{extraembryonic, spermatogenesis, homeobox 1 (Esx1)} (shown), gap junction protein, beta 2 (Gjb2)] are absent or nearly absent from this sample. Boxes, exons; arrows, introns.
Mitosis-promoting genes are down-regulated in endocycling cells. Illumina RNA Sequencing or microarray gene expression levels from murine TGCs (this study), human MKs (1), polytene Drosophila salivary glands (SGs) (2, 3), mouse E14.5 fetal heads (4), and Kc167 cells (5). Shown are boxplots of log₂-transformed quantile-normalized expression ratios of mitotic genes between diploid and polyploid/polytene cells of the same species or between orthologous pairs of mammalian and Drosophila genes. Genes involved in M phase are color coded [purple, polo/polo-like kinase 1 (Plk1); green, cell division cycle 16 (cdc16), fizzy (fzy)/cell division cycle 20 (CDC20), cell division cycle 27 (Cdc27), morula (mr)/anaphase promoting complex 2 (APC2); black, jpl1-aurora-like kinase (jpl1)/aurora kinase B (AURKB), budding uninhibited by benzimidazole 1 (Bub1), mitotic arrest deficient 2 (mad2); blue, twine/string/cell division cycle 25 (Cdc25); orange, cyclin-dependent kinase 4 (Cdk4), cell division control 2 (cdc2)/cycling-dependent kinase 1 (CDK1); red, mitotic cyclins]. Statistical analysis indicates that expression of this set of genes is significantly decreased (one-way ANOVA with a Tukey post hoc test: \( P = 0.008581 \) (TGC/embryo vs. MK16C/2C); \( P = 0.001533 \) (SG/Kc167 vs. MK16C/2C); \( P = 0.000765 \) (SG/brain and discs vs. MK16C/2C) in cells that polyploidize via the endo cycle including the murine TGCs and Drosophila SGs relative to diploid cells. In contrast, >8C MKs, which become polyploid via endomitosis, show the same high expression of mitotic genes as do diploid 2C MKs.


Table S1. List of the 30 most abundantly expressed protein-coding genes in TGCs

Table S2. Gene Ontology (GO) enrichment list and \( P \) values by categories (GO cellular component, GO biological function, GO biological process terms)

Table S3. List of 30 most abundantly expressed miRNAs in TGCs

Dataset S1. Log₂-transformed expression ratios of selected genes shown as colored points in Fig. 4 and Fig. S5