In animals, primordial germ cells (PGCs) give rise to the germ lines, the cell lineages that produce sperm and eggs. PGCs form in embryogenesis, typically by one of two modes: a likely ancestral mode wherein germ cells are induced during embryogenesis by cell–cell signaling (induction) or a derived mechanism whereby germ cells are specified by using germ plasm—that is, maternally specified germ-line determinants (inheritance). The causes of the shift to germ plasm for PGC specification in some animal clades remain largely unknown, but its repeated convergent evolution raises the question of whether it may result from or confer an innate selective advantage. It has been hypothesized that the acquisition of germ plasm confers enhanced evolvability, resulting from the release of selective constraint on somatic gene networks in embryogenesis, thus leading to acceleration of an organism’s protein-sequence evolution, particularly for genes expressed at early developmental stages, and resulting in high speciation rates on zygotic gene activity to specify germ cells at a relatively later stage (1) (Fig. 1) (see SI Appendix, section 1 on the less parsimonious scenario that induction is the derived mode). Induction has been inferred based on microscopic data from most animal clades studied to date and has been experimentally shown to occur in a diverse range of organisms, including mammals (Mus musculus) (4–8), salamanders (Ambystoma mexicanum) (9, 10), and insects such as crickets (Gryllus bimaculatus) (11, 12) and stick insects (Carausius morosus) (13, 14). In turn, inheritance has evolved, apparently independently, in diverse taxa, including nematodes (Caenorhabditis elegans), insects (Drosophila melanogaster and Nasonia vitripennis), cartilaginous fish (Danio rerio), and frogs (Xenopus laevis) (reviewed in ref. 1). In all metazoans, PGC specification occurs relatively early in embryogenesis, but is initiated by mechanisms that differ in their developmental, genetic, and molecular basis, and in the degree of dependence on maternal vs. zygotic genome activity (reviewed in refs. 2, 15, and 16).

To illustrate the major differences between PGC-specification modes, we will briefly describe two well-understood exemplars of each mode: inheritance in the fruit fly D. melanogaster and induction in the mouse M. musculus. In D. melanogaster, PGCs depend critically on a cytoplasmic assemblage of specific, maternally provided gene products (collectively called germ plasm), which is asymmetrically localized to the posterior of the oocyte cytoplasm. During early embryogenesis, cells that form at the posterior pole inherit this germ plasm and adopt PGC fate as a result. A crucial gene for germ plasm assembly is oskar, which is needed to recruit most other germ-plasm components, including, for instance, the nanos transcript, whose protein product helps silence transcription in pole cells and supports the correct migration of PGCs from the posterior pole of the embryo to the mesodermal precursors of the somatic gonad (reviewed in refs. 1, 2, and 15). In summary, the critical feature of the inheritance mode is its reliance on maternally provided germ-line determinants that are asymmetrically localized to the ooplasm and/or early embryonic cytoplasm.

In contrast, the induction mode in the mouse M. musculus does not depend on maternal contributions and is instead reliant on zygotic gene activity to specify germ cells at a relatively later developmental stage (shortly before gastrulation) than they are formed in the fruit fly (well before blastoderm formation). The PGCs are specified by cell–cell signals from the embryonic and extraembryonic endoderm to mesodermal cells of the proximal epiblast. These signals include ligands of the bone morphogenetic protein (BMP) family and the canonical WNT/β-catenin signaling pathways (reviewed in refs. 15 and 16).

Although the upstream gene regulators and developmental processes differ markedly between PGC-specification modes, a number of properties are relatively conserved across taxa with induction and inheritance. For example, expression of many of the downstream effector genes involved in PGC formation and subsequent germ-line development are often conserved across animals, including nanos, vasa, tudor, and piwi (2). Furthermore,

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “Gene Regulatory Networks and Network Models in Development and Evolution,” held April 12–14, 2016, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/Gene_Regulatory_Networks.

Author contributions: C.A.W. and C.G.E. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. Email: extavour@eob.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1610600114/-/DCSupplemental.
PGCs specified under both modes effectively suppress somatic fate (17, 18), presumably to maintain their fate as PGCs. Although the downstream molecular mechanisms underlying PGC fate, expression, and maintenance appear somewhat conserved, it remains unknown why the upstream PGC-specification mechanisms have repeatedly shifted between induction and inheritance among metazoans, and whether these shifts have significant evolutionary consequences.

Here, we consider hypotheses from the literature and present additional proposals, including supporting and refuting evidence, pertaining to the causes and evolutionary consequences of PGC-specification mode in metazoans. An abbreviated summary of the hypotheses discussed is provided in SI Appendix, Table S1.

**Evaluation of the PGC-Specification Hypothesis**

Numerous hypotheses that might explain the putative evolutionary advantage of specifying germ cells early in development have been proposed over the years (discussed in ref. 19), but most of these hypotheses have not been tested empirically or specifically investigated regarding different potential fitness advantages of inheritance vs. induction. To our knowledge, the only such hypothesis that has been subject to explicit testing is what we refer to herein as the “PGC-specification hypothesis” (presented in ref. 3 and 20). This hypothesis asserts that the selective advantage of inheritance is that the early disentanglement of germ-line specification from somatic influences results in the “liberation” of selective pressures on somatic gene regulatory networks (GRNs), thus enhancing species evolvability. In turn, this liberation accelerates an organism’s protein sequence evolution under germ plasm, particularly of genes expressed during early stages of development, and leads to species radiations.

The hypothesis’ predictions pertaining to molecular evolution of genes has been purported to be empirically supported by findings of faster evolution of protein sequences in clades with inheritance than in clades with induction (faster rates were inferred for up to 32% of the genes in a genome under inheritance) based on analysis of four pairs of vertebrates, with one member of the pair exhibiting inheritance and the other induction [anurans vs. urodeles, birds vs. crocodiles/turtles, snakes vs. lizards, and one clade of ray-finned fishes (Teleostei) vs. another (Acipenseriformes) (20)]. As with any hypothesis, testing the predictions by using different methods and systems would be helpful to assessing its generalizability, particularly because that pioneering assessment had some limitations. For instance, because of extreme divergence of taxa being compared and saturation of substitutions, only nonsynonymous coding-DNA substitutions (dN) were studied, excluding synonymous site changes (dS) and dN/dS, which would have been needed to assess or make conclusions about variation (or liberation) in selective pressures under preformation (21, 22). Importantly, the study comprised numerous overlapping nonphylogenetically independent contrasts, which is pseudoreplication (23, 24). Small gene sample sizes for some taxa, the use of tissue-specific transcriptomes, and exclusion of invertebrates, which comprise the vast majority of animal life, also may have limited those findings (25) (see further details on these limitations in SI Appendix, section 2). Accordingly, further investigations using alternate approaches and organisms would be useful to assess whether the purported liberation of constraint and increase in rates of protein-coding sequence evolution under germ plasm (20) is robust to different types of methods and animal systems.

Our recent comparative large-scale molecular evolutionary analysis using a different study approach and different datasets, and spanning both vertebrates and invertebrates, did not support the predictions of the PGC-specification hypothesis (25). The assessment was based on genome-wide dN/dS of phylogenetically independent species pairs from 12 genera with different PGC-specification modes [the invertebrate genera *Drosophila, Nasonia, Schistosoma, Anopheles, and Pristionchus* (inheritance) and *Trichoplax, Echinococcus, and Apis* (induction); and the vertebrate genera *Falco* and *Xenopus* (inheritance), and *Alligator* and *Pan* (induction)]. The results of these analyses supported the null hypothesis that PGC-specification mode has no consistent effect on protein sequence evolution, including on the evolution of the sequences of early developmental genes (25). Thus, these recent findings suggest that, at a minimum, the PGC-specification hypothesis does not hold in some animal genera, and that there is a good possibility that germ plasm is not strongly, or potentially even marginally, linked to rapid protein sequence divergence. Although further testing in even more taxa will be helpful as more genomic data become available, these results already comprise significant examples countering the notion that germ plasm generally and broadly releases selective constraint on genes and somatic gene networks in animals, suggesting that alternate causes of the evolution of germ plasm should also be explored.

**Inheritance Mode May Not Be Linked to Speciation.** An additional prediction of the PGC specification hypothesis is that the disentanglement of the germ line and the soma during early embryogenesis should allow freedom for morphology and protein sequences to evolve, and thus lead to species radiations greater than those typically observed under induction (3, 20, 26). Extreme differences in species richness reported between some vertebrate clades have been taken as support for this prediction. For example, some clades that specify germ cells using inheritance have higher reported species numbers (e.g., frogs,
4,800 species; teleosts, 25,000; birds, 10,000; and ascidians, 3,000) than related taxa that use induction (e.g., salamanders, 515; other bony fishes, 44; turtles, 300; and hemichordates, 100) (species numbers are cited in ref. 3). A later study used a similar approach to provide examples of greater species richness in metazoan clades with induction. In one invertebrate example, the Mollusca, groups with germ plasm (summed across Cephalopoda, Gastropoda, and Bivalvia) were reported to have 90,900 species, whereas those with induction (Aplocaphora, Monoplocaphora, and Polyplacophora) had 3,195 (26). These studies, based on reported species richness alone, have concluded that the observed trends support the prediction that inheritance mode provides a release of selective constraint, enhancing evolvability, and results in increased rates of speciation in animals (3, 26).

It is also worth considering, however, that species richness alone can be a poor indicator of speciation, or diversification rates, because it does not include the role of clade age, birth–death rates, or any molecular evolutionary sequence analyses, absences that might lead to misleading conclusions (27, 28). We provide examples in SI Appendix, section 3 suggesting that germ plasm might not be a causative factor strongly or typically linked to high speciation rates, at least in some animal groups. For example, no effect of preformation is detectable when taking into account diversification-rate data in frogs and salamanders (29). Various other jawed vertebrates, including both preformation taxa (birds and teleosts) and induction taxa (lizards and eutherian mammals), each exhibit accelerated (nontypical) diversification rates (29). Furthermore, species richness itself appears unconnected to PGC-specification mode in multiple invertebrate lineages (25). Collectively, such examples appear inconsistent with the PGC-specification hypothesis (25).

A final crucial aspect of the prediction of enhanced speciation under germ plasm (3, 20) that warrants consideration is the mechanism of reproductive isolation. For instance, the notion that greater speciation occurs under germ plasm would require a proposed mechanism for enhanced instances of reproductive isolation, which give rise to speciation events, under this PGC-specification mode. We speculate on suitable arguments one could use to explain reproductive isolation under preformation in SI Appendix, section 3, which include rapid divergence of male–female fertilization genes (cf. ref. 30) or the somatic recombination of GRNs (with Clade age, birth–death rates, and any molecular evolutionary sequence analyses, absences that might lead to misleading conclusions (27, 28). We provide examples in SI Appendix, section 3 suggesting that germ plasm might not be a causative factor strongly or typically linked to high speciation rates, at least in some animal groups. For example, no effect of preformation is detectable when taking into account diversification-rate data in frogs and salamanders (29). Various other jawed vertebrates, including both preformation taxa (birds and teleosts) and induction taxa (lizards and eutherian mammals), each exhibit accelerated (nontypical) diversification rates (29). Furthermore, species richness itself appears unconnected to PGC-specification mode in multiple invertebrate lineages (25). Collectively, such examples appear consistent with the PGC-specification hypothesis (25).

A Related Hypothesis on PGC Specification. It should be noted that a related hypothesis on the evolution of PGC-specification mechanisms has been proposed based on the developmental timing of PGC formation. In particular, Johnson and Alberio (32) recently presented a hypothesis (denoted hereafter as the deterministic-stochastic hypothesis) addressing the "timing and mechanisms of PGC specification in the vertebrate lineage." In this report, it was proposed that the developmental timing of PGC establishment during embryogenesis, rather than the molecular mechanism (inheritance or induction) per se, underlies liberation of selective constraint on the evolution of GRNs for somatic development, and thus drives species evolvability. Under this hypothesis, PGC-specification modes are classified either as deterministic (early PGC specification; that is, effectively those taxa with germ plasm, as well as some induction taxa with "early" PGC formation, specified therein as rodents) or stochastic (late specification, presumably including many or most induction taxa). Because mouse (induction) uses the transcription factor Blimp-1 early in embryogenesis (day 6.25), which commits cells to the germ-line fate after separation of embryonic from extraembryonic tissues, but before specification of major groups of embryonic somatic lineages (shortly before gastrulation), it was argued this taxon exhibits a "deterministic" mode of PGC formation. In contrast, species like salamanders exhibit a "stochastic" mode of PGC specification and form PGCs late in embryogenesis (germ-line commitment occurs after gastrulation); in this case, PGCs are the last cells in the embryo to engage in lineage commitment, a phenomenon described as the "last cell standing model." Thus, based on this premise, mice, despite using induction to specify PGCs, should nevertheless exhibit release of selective constraint on somatic GRNs, rapid evolution (enhanced evolvability), and high rates of speciation, similar to what was predicted for organisms with germ plasm, whereas others, such as salamanders, evolve slowly (3, 20).

The predicted trends in mouse genome evolution, including enhanced speciation, under the deterministic-stochastic hypothesis were suggested to be consistent with their high species richness (of the reported 2,277 species of rodents, ~61% are Muridae (32)). It was also suggested that the faster evolution reported for mouse gene sequences (based on analysis of substitutions of amino acids and each of three codon positions for three protein-coding genes, and of a ribosomal RNA; ref. 33) compared with other mammals concurs with the deterministic-stochastic hypothesis (32). However, rapid sequence evolution in the mouse lineage is well known, and need not indicate a release of selection, but, rather, is likely an effect of their short generation time (34, 35). As an example, an assessment of selection using dN/dS values (988 genes) has shown that mice (M. musculus) exhibit lower average values than other mammals such as pigs (Sus scrofa) and humans (Homo sapiens), suggesting greater purifying selection, and thus inconsistent with the extensive release of constraint predicted by the deterministic-stochastic hypothesis (ref. 34; see also ref. 35). Mice may be subjected to more restrained evolution than some other mammals simply due to their very large population sizes (34). Furthermore, similar to the PGC-specification hypothesis, this hypothesis contends that inheritance accelerates evolution (early PGC specification, classified as deterministic), and thus is discordant with examples suggesting no such effect in some animal lineages (induction taxa, presumably classed as stochastic; ref. 25). Nonetheless, whether ultimately well supported or disproven by future studies, the hypothesis highlights that the timing of induction varies among taxa with this PGC-specification mode, and that the stage of inductive signaling might be biologically or evolutionarily meaningful.

This more recent deterministic-stochastic hypothesis (32) is related, but differs in some contexts, from the prior PGC-specification hypothesis, which asserted that inductive signaling in embryogenesis acts as a constraint and slows animal evolution, without respect to timing of PGC specification (3, 20). In particular, the more recent hypothesis regarding embryonic timing of lineage commitment (32) appears to suggest that the prior PGC-specification hypothesis (3, 20) should be modified, such that slowed evolution under induction (compared with inheritance) would only be predicted in those taxa where the inductive signaling occurs late in embryogenesis. To further assess the more recent deterministic-stochastic hypothesis, studies should quantify PGC specification in induction taxa at early (preferably extending beyond rodents) vs. late embryonic stages, or relative to the specification of various somatic lineages. It will be a challenging, but necessary, aspect of such an analysis to determine usefully comparable demarcations between early and late stages across taxa, and clear definitions of “lineage commitment” beyond the level of germ layers, to ascertain whether
any consistent effect on molecular evolutionary rates, speciation, and/or embryo development are observed with respect to timing of PGC formation during ontogeny within induction taxa.

**Evolution of Inheritance May Not Necessarily Depend on its Effect on the Soma**

Because previous and detailed arguments have already been made for why germ plasm should accelerate evolution (3, 20), here we consider the other possible scenario—that is, that the null hypothesis may hold for many or most animals, and that germ plasm does not give rise to rapid animal evolution, including a broad effect on gene sequence evolution (see SI Appendix, section 4 for discussion of a possible smaller-scale effect on the evolution of specific genes). To explain the phylogenetic distribution of PGC-specification modes across animals, and extended to the PGC-specification hypothesis, it was proposed that “the distribution of epigenesis and preformation must result from the influence each mode of germ cell specification has on the development of the soma” (3). The proposition that inheritance (preformation) must evolve convergently because of its effect on somatic tissues was suggested to be supported by certain characteristics of the genetic mechanisms regulating somatic development. For example, it was noted that frogs and teleosts, which specify germ cells by inheritance and exhibit a complex GRN governing embryonic mesoderm formation, possess multiple copies of the mesoderm inducers Nodal and Mix. In contrast, Axolotls (Ambystoma mexicanum, Mexican salamanders), which specify germ cells by induction, contain just one copy of each of Nodal and Mix. This finding was interpreted as resulting from greater constraint in the induction species and the evolution of novelty in the frog GRN (and presumably also in teleost GRNs) due to the liberation of constraint under germ plasm (ref. 3; see also ref. 32). Although this scenario indeed comprises one feasible possibility, it is also worth considering that the differences in mesoderm GRNs between these taxa for this sample of genes might result from various other factors. For instance, it is feasible that there is a greater tendency for tandem duplications of chromosomal regions, neo-functionalization and sub-functionalization, and/or fewer (duplicate) gene losses in frogs and teleosts (36–38), factors potentially unconnected to PGC-specification. Another contributing factor may be whole-genome duplications, which have been linked to rapid evolution of duplicated genes and are believed to have occurred within some frog (e.g., X. laevis; ref. 39) and teleost lineages (36, 40, 41), but likely not in the salamander A. mexicanum (42). In this regard, although fewer gene copies under induction could reflect greater constraint on these gene pathways (3), further studies should aim to further disentangle such an effect from other plausible factors unrelated to PGC-specification mode. For example, it would be useful to evaluate patterns of duplications of developmental patterning genes, their evolution, and their distribution among inheritance vs. induction species across a wide range of animal taxa.

**Alternate Scenarios Possibly Explaining the Evolution of Germ Plasm.**

The inference that the evolution of germ plasm must result from its effect on, and release of natural selection within, the soma (3, 20) may not be strongly supported in premise based on available literature. An argument can be made, for example, that somatic selective pressures do not need to be operative and that germ plasm might have arisen via other, largely somatically unrelated, evolutionary forces. For instance, germ plasm need not necessarily arise because of selective effects operating on the soma, but rather, could evolve from selection within the PGCs and/or the germ lines to which they give rise. Selection among cells within an individual, which studies show can include cell-lineage selection within the germ lines, has been suggested to be a relevant mechanism in animal evolution (43, 44) and may lead to preferential differential transfer of certain mutations (or allele combinations) to the offspring (43–46). Specifically, natural selection between cells with different cellular phenotypes and their causal alleles in the PGCs and germ lines provides an avenue for preselection of mutations within the germ line before their effects are manifest in the soma. In other words, intragerm line selection could contribute to the removal of deleterious germ-line mutations and the promotion of beneficial ones to later generations (43). Such mutations may or may not be favorable for the soma, but their differential inheritance would be based largely on their phenotypic effect on the germ line, rather than the soma (43, 45, 46).

Furthermore, during and after meiosis, in gametogenic cells and gametes, even recessive haploid mutations, those typically sheltered by diploidy, may be subjected to selection (30, 47, 48) and possibly contribute to the evolution of germ plasm. For instance, because germ plasm is present in eggs and/or zygotes/early embryos, it is plausible that sexual selection may play a role: Sexual antagonism between genes involved in egg–sperm interaction during fertilization or male-mate choice affecting egg traits (such as germ plasm or correlated female traits) may contribute toward the evolution of the inheritance mode (30, 48). In this respect, maternal and/or gene expression changes that lead to the switch from induction to inheritance could potentially be fixed by cell-lineage selection in the precursors of the PGCs, the PGCs themselves, or the germ-line cells and/or sexual-selection pressures on the sex cells. It is also possible that germ plasm arises from an effect on a small component of somatic genes specifically involved in cells giving rise to PGCs. Accordingly, the distribution of induction and inheritance across animals need not, in principle (3, 20), be driven by their broad influence on the soma or somatic GRNs (25).

Although germ plasm may arise in response to selection, it is also possible that germ plasm is a side effect, or spandrel (49, 50), rather than a direct target of selection. In other words, it may result from indirect selection or be a by-product of selection on another connected biological feature (49, 50). For instance, organisms displaying the inheritance mode also tend to heavily rely on maternal determinants for early axial patterning and body plan specification, and these determinants are often formed by asymmetric deposition of molecules within the oocyte during oogenesis and early embryogenesis (51). Along with these determinants for somatic patterning, germ-line determinants are often included in the battery of asymmetrically localized molecules. If convergent shifts toward body patterning shaped by maternal determinants become advantageous in different taxa, then it may be inevitable for these taxa to acquire germ plasm via similar maternally derived mechanisms as part of a streamlined system of development. In this respect, germ plasm may be a by-product of a system shifting away from regulative embryonic development and toward increased use of maternal determinants of body plans. Drosophila (fruit flies), Danio (zebrafish), and Xenopus (frogs), which are among the most prevalent laboratory models used in developmental biology, all use maternal determinants to direct body patterning in embryos (52–54) and also use germ plasm for PGC specification (reviewed in ref. 1), consistent with the involvement of maternally derived RNAs and/or proteins in both developmental processes. In M. musculus (mice), which forms PGCs by induction (reviewed in ref. 1), the idea that localized maternal determinants are used in embryonic axial patterning remains controversial. For example, some studies, but not others, find evidence for a maternal role of gene products proposed to be maternal determinants (Dcd2) (55, 56) in cell-fate specification in the embryo (57–59). Nevertheless, a significant body of data suggests that zygotic regulators and cell–cell interactions determine axis polarity and patterning in mammals (e.g., refs. 60–63). Thus, together, these inheritance and induction models agree with a correlation between germ plasm and maternal determination of axial patterning, and a spandrel effect. Nonetheless,
these putative trends could simply be an artifact of the particular model systems studied to date. We suggest such an artifact as a possibility for consideration, given that a majority of animal model organisms in developmental biology exhibit relatively rapid life cycles, highly stereotypical (canalized) development, tolerance of high population densities, and variable environmental conditions. These are the qualities that make them convenient and manipulatable organisms for laboratory study. However, prioritizing these features can also yield to choices of model organisms that display some developmental similarities to each other, but are not representative of the larger taxa to which they belong. Furthermore, some cnidarians and echinoderms use induction for PGC specification (1) and yet certain data have suggested that maternal products deposited in the egg (products of the Fruzzled protein family that activate Wnt signaling in cnidarians and Panda gene products in echinoderms) may partly contribute toward directing embryonic body plans in these groups, raising the possibility that the mechanism of PGC formation and axial patterning are uncoupled in some organisms (64, 65). Together, further investigations of mechanisms of embryonic axis formation and PGC specification modes across a wider range of metazoans will be needed to decipher whether a shift toward use of maternal determinants for embryonic patterning typically cooccurs with germ plasm, consistent with a spandrel effect.

Is the Transition to Inheritance Mode Irreversible and Convergent?

The distribution of inheritance and induction modes across metazoans (Fig. 1) suggests that induction is ancestral to Bilateria and prevalent throughout both protostomes and deuterostomes, whereas inheritance has been derived in multiple lineages (1, 19). Although there are numerous examples suggesting that the inheritance mode arises from an ancestral induction mechanism (1), reports of clades that might exemplify a transition from induction to inheritance remain rare (1, 66), suggesting that the shift to germ plasm is typically irreversible. Germ plasm thus appears to follow “Dollo’s Law”—that is, structures or processes lost in evolution are unlikely to be regained by descendants in the same form as the ancestors (67, 68). In this aspect, the inheritance mode resembles other typically irreversible transitions, such as the transition from outcropping to selfing and from hermaphroditism to dioecy (69). In placental mammals, for example, the loss of a gene encoding an enzyme in the anthocyanin pathway is needed for the transition, which limits the opportunity for later reversal to the ancestral state (69, 70). Accordingly, the transition from induction to inheritance could, in principle, be made irreversible by a single gene mutation or gene loss in the induction pathway. Although it is known that germ plasm has a distinct molecular genetic mechanism from that of induction, as observed in the contrast between the mechanisms used by the models D. melanogaster and M. musculus (2), this difference alone should not be sufficient to prevent a reversal to induction, unless the induction pathway has been impaired by within-gene mutations, gene losses, or gene silencing. Indeed, although key genes needed for induction in mice and crickets, such as BMP signaling pathway members and specific downstream transcription factors (2, 71), are highly pleiotropic, they show no signs of being used in early PGC specification in Drosophila (but see ref. 72 for evidence of a role for BMP signaling in PGC fate maintenance), or in other less-well-studied taxa with germ plasm (2). Furthermore, the ablation of PGCs or their precursors in animal models with inheritance (e.g., D. melanogaster, C. elegans, D. rerio, and X. laevis) is generally not corrected by inductive signaling and de novo establishment of PGCs, suggesting that a putative ancestral inductive PGC-specification mechanism has been lost in those taxa (see, for example, ref. 73). In this regard, gene loss, and possibly mutations or silencing in upstream regulators under induction, could contribute to an irreversible transition to the inheritance mode of PGC specification in metazoans. Nonetheless, during or shortly after the transition from induction to inheritance, one might expect gradual loss of induction mechanisms (19), and thus in this period a taxon could exhibit some reversal capabilities, as has been implied for the solitary ascidian Ciona intestinalis (74). Further studies will be needed to ascertain whether unambiguous examples of species exhibiting both PGC-specification modes can be identified, in support of this hypothesized transition period, or whether a complete irreversible transition to germ plasm, including the loss of induction functionality, is prevalent and likely occurs rapidly.

Under a presumption that the inheritance mode comprises an adaptation, one must question whether this adaptation occurs via divergent or similar genetic mechanisms. Unlike other derived traits—such as the transition to wings in insects, birds, and bats (75) or from a primitive photoreceptor to the camera-eye found in octopus and vertebrates (76, 77), both of which have apparently unambiguous advantages in terms of adapting to environmental conditions—an adaptive advantage of germ plasm appears to be less obvious. Convergent phenotypes can arise from independent genetic pathways, as is believed to occur for wing formation (75, 78), or can result from similar genetic mutations that arise in independent lineages. The latter may include convergent phenotypes arising from a mutation in orthologs or orthologous pathways between lineages (parallel evolution), a shared allele that was polymorphic in ancestral populations, or from shared introgressions (collateral evolution) (75). In the case of germ plasm, the data to date suggest that this convergent phenotype results from distinct genetic pathways in different lineages. For example, the oskar gene has been shown to be sufficient for germ-plasm assembly in D. melanogaster (79–81). This gene has been identified in a number of insect lineages (66, 82), but lacks known orthologs in noninsect metazoan lineages, including those with germ plasm (e.g., birds, fish, and frogs) (82). The novelty of oskar to insects is one indication that germ plasm arose independently via different genetic mechanisms across the metazoans, at least between the insects and noninsect animal systems. Consistent with the notion of independent genetic regulators of germ plasm across animals, the bucky ball gene in the zebrafish D. rerio has been shown to be required for germ-plasm assembly in that taxon (83–85), whereas this process is believed to be modulated by the nematode-specific MEG (maternal-effect germ-cell defective) and PGL (P-granule abnormality) genes in the model nematode C. elegans (86, 87). It is worth noting that oskar (Drosophila), bucky ball (Dario), and MEG/PGL (Caenorhabditis) differ in their pleiotropic roles. For example, although all genes are involved in germ-plasm function or assembly, only the former two genes have been shown to also play a role in embryo or oocyte axial polarity (53, 85, 88–90), findings consistent with differences in their evolutionary dynamics among taxa. In summary, it appears feasible that at least in some lineages, inheritance arose rapidly, potentially due to rare highly beneficial mutations, novel gene evolution, and/or introgressions across populations, leading to fixation of germ plasm based on only a few genetic changes (75, 91).

Given that no specific example of shared mutations or introgressions have yet been identified that could explain every instance of the evolution of germ plasm across metazoans, the most plausible scenario is that this phenotype arose repeatedly via different genetic mechanisms. Furthermore, the biological properties of germ plasm differ in some respects across taxa: Germ plasm can originate as a molecular accumulation in the oocytes before fertilization (D. melanogaster, D. rerio, Gallus gallus, and X. laevis) or clustering and concentration of P granules in the early embryos (C. elegans), and each of these mechanisms will ultimately drive formation of PGCs in early embryogenesis (92).
In addition, a number of marked differences have been noted between *Xenopus* (frogs) and *Danio* (zebrafish) species in terms of the embryological location and specific molecular components of germ plasm (3). Thus, germ plasm appears to be functionally (or analogously) convergent, but the underlying mechanisms are not necessarily perfectly developmentally or genetically convergent. Together, these findings concur with the hypothesis of convergent evolution of inheritance as a PGC-specification mechanism. An argument is therefore available for the existence of an innate selective advantage to germ-line determination via localized maternal germ plasm, causing it to arise independently, with highly similar functionality, but with different genetic mechanisms and biological properties, across diverse systems.

**Reproductive Lifestyles and Germ Plasm**

A fitness advantage of germ plasm compared with induction is not currently known. Speculatively, however, it can be considered that germ plasm might be particularly beneficial under a change to an unfavorable or dynamic environment (referred to hereafter as “stressful”) that reduces reproductive output or germ-line and embryo survival. In principle, when the maternal determinants are synthesized within the female sex organs, it may make the biochemical cost of development in the early embryo substantially lower, by reducing costs of RNA and protein synthesis and/or cellular transport to asymmetrically localize axial or regional determinants, thus increasing germ-line establishment and embryo survival rates under stress. For oviparous organisms (egg-bearing reproduction), germ plasm may be particularly beneficial, because the molecules would not need to be synthesized in a nonmaternally supported, independent embryo under a shift in environmental conditions, as they would under the induction mode. In viviparous (and ooviviparous; live-bearing reproduction) organisms, where the embryo remains supported by the female throughout embryogenesis and PGC formation, a reduced cost in establishment of the germ line and embryo proper could also, in principle, be advantageous to offspring survival. Nonetheless, assuming germ plasm is beneficial to survival under a dynamic changing environment, one may speculate that any putative benefit of germ plasm might be particularly elevated for oviparous organisms that lack the ongoing support/investment from maternal tissues, which can reduce the production and viability of offspring (93). This is further discussed in *SI Appendix*, section 5, reproductive lifestyles might be a significant factor influencing the evolution of germ plasm in animals.

**Germ-Line Segregation and Mutation Rates**

Given that the PGCs and germ lines are the source of all heritable genetic mutations, and thus genetic variation available in biological systems, they play a central role in evolution. Mutations are thought to arise primarily from DNA-replication errors, but can also result from faulty DNA repair, transcription-mediated mutation, and environmental and physiological agents (47, 94–98). The mutation rate in coding DNA per generation underlies a diverse set of evolutionary phenomena, including the evolution of sex, aging, recombination, mating systems, species extinctions, reproductive isolation, and speciation (30, 99). Evidence to date has shown that the mutation rate in coding DNA varies among animals (95, 100, 101), and the rate itself can be subject to selective pressures (98, 101, 102). Given the central role of germ-line mutations to evolutionary biology, it is worth considering whether and how these mutations—and particularly mutation rates in the germ lines—could be related to the evolution of the germ-line soma divide and to PGC-specification mode in animals. We highlight putative differences in the mutation rates of the germ line and soma in *SI Appendix*, section 6, and describe how PGC-specification mode may affect germ-line mutation rates below.

**How PGC-Specification Mode Could Influence Germ-Line Mutations.**

Males typically have higher germ-line mutation rates per generation than females (48). This phenomenon is believed to be predominantly due to the higher number of cell divisions, and thus replication errors, that take place in male germ lines (47, 48, 103, 104), but may also partly depend on other variables (e.g., methylation; ref. 47). In humans, analysis of mutations in multisolviding families has shown that the PGC-specification (de novo) mutation rate per cell division (~10 cell divisions before PGCs are formed) was 0.2–0.6 mutations per haploid genome per cell division for both maternal (includes oogenesis) and paternal germ lines and was 0.5–0.7 for the post-PGC stage to puberty (~20–24 post-PGC cell divisions up to puberty). Thus, not only the post-PGC stage, but also the pre-PGC stage appears to be a significant factor contributing toward the genomic mutation rate. Notably, the postpuberty mutation rate per cell division (0.09–0.17) in males was markedly lower than prepuberty (lower than pre- and post-PGC), trends believed to result from selection to reduce the mutation rate to compensate for the high number of cell divisions involved in sperm formation (105). The male-to-female ratio of de novo mutations in offspring occurred at a 3.5:1 ratio, approximately corresponding to values reported previously for humans and consistent with the higher number of male (than female) germ-line cell divisions per generation (105, 106). However, the de novo mutations specifically arising within the cells before PGC specification had a 1:1 ratio of maternal and paternal origin, as would be expected for mutations that arose before PGC formation and male/female differentiation (105). Together, these data empirically support the notion that the pre-PGC stage (before separation of the germ line and soma) can make a significant contribution toward the mutation rate per generation (105, 107). Although the pre-PGC-specification mutations occur in a small fraction of embryonic cells, and arise in males and females, their early origin suggests that they may occur in a majority of germ-line cells and in the next generations, depending on cell-cell selection. Given that the two distinct PGC-specification modes (inheritance and induction) vary developmentally and genetically at the pre-PGC stage (1, 2, 15), the mode could have a significant influence on whether and how many pre-PGC mutations are transferred to the germ lines and thereby to the offspring. Further studies of the genome dynamics at pre-PGC, as well as at post-PGC stages, using additional animal models representing both modes of PGC specification (and models with variable biological properties within inheritance mode), will be needed to ascertain how PGC specification mechanisms may influence mutation rates and cell-cell selection during these early developmental stages.

**PGC Specification May Be Linked to the Germ-Line Mutation Rate.**

Inheritance and induction may differentially affect mutation rates. The length, number of cell divisions, and cellular behaviors during pre-PGC and PGC stages differ between induction and inheritance modes (2, 15, 19), and thus the rate of mutation might vary among modes. Furthermore, under inheritance, the PGCs typically form earlier (blastoderm stage) in embryogenesis than under induction (gastrulation), and thus may be mitotically and transcriptionally quiescent for an extended period after their specification (15, 17, 18, 92), possibly favoring a lower mutation rate per generation. Some available mutation rate data indicate that the induction taxa mice and humans exhibit a higher germ-line mutation rate (~38.00 (note: male germ line) and 12.85 × 10^{-9} per site per generation, respectively) (101) than the inheritance vertebrates *C. elegans* (5.60 × 10^{-9}) (101) and *D. melanogaster* (4.65 × 10^{-9}) (101, 108, 109) [also see other factors, such as population size (101)]. Recent direct estimates of mutation rates using whole-genome DNA sequencing indicate a lower rate in the bird *Ficedula albicollis* (inheritance) than humans and chimpanzees (induction) (see ref. 110 for additional taxa such as mouse). A lower mutation rate may be selectively advantageous (101, 102), and thus...
speculatively might sometimes contribute toward the evolution of inheritance mode in animals.

At present, however, there are insufficient available studies, and a wider range of direct mutation-rate data (110) from animals with known PGC-specification modes will be necessary to ascertain any effect on mutation rates. Furthermore, even if PGC-specification mode shapes mutation rates before or during PGC specification, depending on the strength of effect, it might not be strongly correlated to mutation rates across species, because this value also depends on numbers of mutations (and cell divisions) arising after PGCs are specified (110). Furthermore, mutation rates may vary with other factors such as metabolic rates (111), individual age (112), and natural selection on mutation rates, which depends on population size (102). Thus, disentangling the role of PGC-specification mode might require multifactorial assessments of mutation rates across animals wherein all these parameters have been well established, to isolate any PGC-specification mode effect. We note that recent dN/dS data suggests that innate mutation rate variation between frog (inheritance) and salamander (induction) lineages likely explains how PGC specification mode per se influences mutation rates in those taxa. To better understand whether PGC-specification mode influences cell–cell selection on mutations in the germ-line lineage in metazoans, future studies should also assess the evolution of genes specifically expressed in pre-PGC (and post-PGC) cells to measure the effects of selection (such as dN/dS, ref. 21) and include laboratory techniques involving measures of fitness effects of specific mutations in those cells and of cell–cell competition (44, 105).

Conclusions

Here, we have described existing hypotheses from the literature and set forth additional hypotheses and proposals for further consideration, with respect to the causes and consequences of PGC specification mechanisms in metazoans (SI Appendix, Table S1). Together, the data to date suggest that the transition to germ plasm in metazoans occurred convergently via different genetic and developmental mechanisms, which may have involved adaptive processes, or, alternatively, may have arisen as a spandrel effect. Furthermore, PGC specification may be connected to life history parameters such as oviparity and viviparity. We argue that, because PGC specification mode is indispensable for germ-line formation, it is apt to affect the germ-line genomic mutation rate, which is one of the most crucial parameters in evolutionary biology. Expanding research in this area will thus be essential to gaining an understanding of the nature of that relationship.

ACKNOWLEDGMENTS. We thank Eric H. Davidson, participants of the Sackler Colloquium on “Gene Regulatory Networks and Network Models in Development and Evolution” for discussion of these issues, and two anonymous reviewers for valuable comments on an earlier version of the manuscript. This work was partially supported by funds from Harvard University and NIH Grant 1 R01 HD073499 (NICHD) (to C.G.E.).

80. Harris WA (1997) Pax-6: Where to be conserved is not conserva