Causes and evolutionary consequences of primordial germ-cell specification mode in metazoans

Carrie A. Whittle[a] and Cassandra G. Extavour[a,b,1]

[a]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; and [b]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

Edited by Neil H. Shubin, The University of Chicago, Chicago, IL, and approved February 1, 2017 (received for review October 12, 2016)

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 in germ plasm-containing lineages (denoted herein as the "PGC-specification hypothesis"). Although that hypothesis, if supported, could have major implications for animal evolution, our recent large-scale coding-sequence analyses from vertebrates and invertebrates provided important examples of genera that do not support the hypothesis of liberated constraint under germ plasm. Here, we consider reasons why germ plasm might be neither a direct target of selection nor causally linked to accelerated animal evolution. We explore alternate scenarios that could explain the repeated evolution of germ plasm and propose potential consequences of the inheritance and induction modes to animal evolutionary biology.

germ line | preformation | epigenesis | primordial germ cell | spandrel

Primordial germ cells (PGCs) are specialized cells located in sexual organs of animals. These cells are central to reproduction because they give rise to the germ lines and, ultimately, the sperm and eggs. The PGCs typically arise during early embryogenesis by one of two distinct modes: (*i*) induction, wherein germ cells are induced by cell–cell signaling pathways (also known as epigenesis); or (*ii*) inheritance, wherein germ cells are formed by using germ plasm (a specialized cytoplasm containing proteins and RNAs needed for PGC formation), that is, maternally generated germ-line determinants that are "preformed" before embryogenesis begins (also known as preformation) (1–3). The induction mode appears to be the more prevalent mode in animals and occurs in basally branching lineages, and thus is the hypothesized ancestral mechanism of PGC specification, whereas inheritance comprises the likely derived state (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.

[1]To whom correspondence should be addressed. Email: extavour@oeb.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1610600114/-/DCSupplemental.

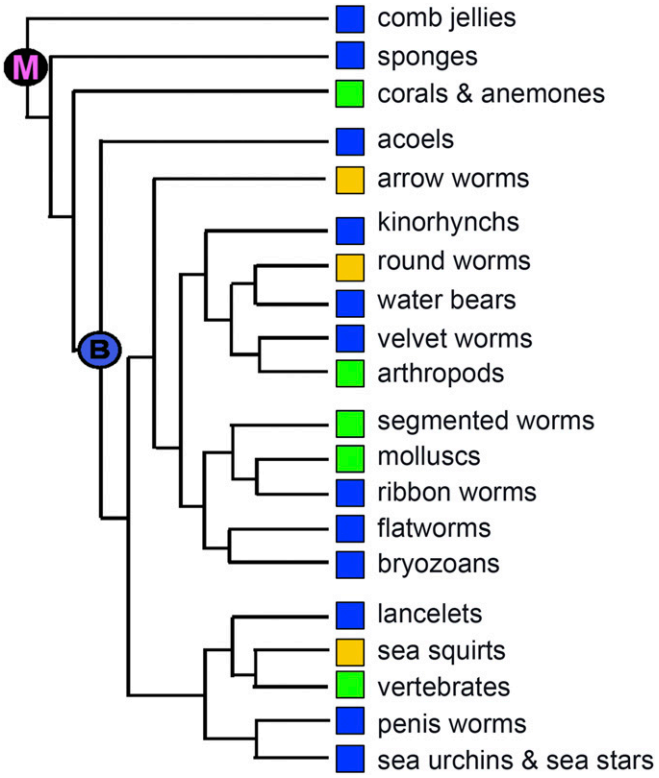

Fig. 1. Phylogenetic distribution of PGC specification mechanisms across Metazoa. Shown is a simplified phylogeny of animals (as per ref. 114), indicating whether members of a given clade exhibit inductive PGC specification (blue) or inheritance-mediated PGC specification (yellow), or whether the clade contains some members reported to use inheritance and others reported to use induction (green). Letters in circles at nodes indicate Bilateria (B) and Metazoa (M).

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), an approach known to cause spurious correlations (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 *Tribolium*, *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 (Aplacophora, Monoplacophora, 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 reproductive organs [and underlying sex- and reproduction-related genes (SSR); ref. 31]. Together, stronger data directly linking germ plasm to speciation rates, combined with some evidence consistent with an underlying mechanism, would be needed to support this particular aspect of the hypothesis.

At present, therefore, the various facets of the PGC-specification hypothesis (3, 20) have multiple examples of genera or clades that refute or counter its predictions, suggesting that it would be prudent to contemplate that factors other than a broad release of selective constraint on the soma and somatic GRNs could explain the repeated evolution of germ plasm.

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 extra-embryonic 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 determinative), 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 (*A. 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, neofunctionalization and subfunctionalization, 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, mutational and/or gene expression changes that lead to the switch from induction to inheritance could potentially be fixed by cell-linage selection in the precursors of the PGCs, the PGCs themselves, or the germ-line cells and/or by 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 (*Cdx2*) (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 *Frizzled* 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 outcrossing to selfing and from hermaphroditism to dioecy (69). In plants, for example, the shift from blue to red flowers in Andean *Iochroma* (Solanaceae) is irreversible; molecular studies suggest that 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 critical 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 egg/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 thus 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 ovoviviparous; 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 an offspring's probability of survival (93). Thus, as 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 multisibling families has shown that the pre-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 invertebrates *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 stages, 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 fast evolution in the former group (113), and not a broad release of constraint in frogs (as dN/dS was similar among taxa), as had been claimed for these taxa (using dN alone) under the PGC-specification hypothesis (20). Direct measures of the mutation rate within a generation (110) and per unit time, combined with multifactorial assessments, will help further discern whether and 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 the manuscript. This work was partially supported by funds from Harvard University and NIH Grant 1 R01 HD073499 (NICHD) (to C.G.E.).

1. Extavour CG, Akam M (2003) Mechanisms of germ cell specification across the metazoans: Epigenesis and preformation. *Development* 130(24):5869–5884.
2. Ewen-Campen B, Schwager EE, Extavour CG (2010) The molecular machinery of germ line specification. *Mol Reprod Dev* 77(1):3–18.
3. Johnson AD, Richardson E, Bachvarova RF, Crother BI (2011) Evolution of the germ line-soma relationship in vertebrate embryos. *Reproduction* 141(3):291–300.
4. Tam PP, Zhou SX (1996) The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev Biol* 178(1):124–132.
5. Ying Y, Qi X, Zhao GQ (2001) Induction of primordial germ cells from murine epiblasts by synergistic action of BMP4 and BMP8B signaling pathways. *Proc Natl Acad Sci USA* 98(14):7858–7862.
6. Ying Y, Zhao GQ (2001) Cooperation of endoderm-derived BMP2 and extraembryonic ectoderm-derived BMP4 in primordial germ cell generation in the mouse. *Dev Biol* 232(2):484–492.
7. Lawson KA, et al. (1999) *Bmp4* is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev* 13(4):424–436.
8. Hogan BL (1996) Bone morphogenetic proteins: Multifunctional regulators of vertebrate development. *Genes Dev* 10(13):1580–1594.
9. Nieuwkoop PD (1947) Experimental observations on the origin and determination of the germ cells, and on the development of the lateral plates and germ ridges in the urodeles. *Arch Neerl Zool* 8(1):1–205.
10. Kotani M (1957) On the formation of the primordial germ cells from the presumptive ectoderm of *Triturus* gastrulae. *J Inst Polytech Osaka City Univ D* 8: 145–159.
11. Donoughe S, et al. (2014) BMP signaling is required for the generation of primordial germ cells in an insect. *Proc Natl Acad Sci USA* 111(11):4133–4138.
12. Ewen-Campen B, Donoughe S, Clarke DN, Extavour CG (2013) Germ cell specification requires zygotic mechanisms rather than germ plasm in a basally branching insect. *Curr Biol* 23(10):835–842.
13. Cavallin M (1976) La ségrégation de la lignée germinale chez le Phasme Carausius morosus Br. *Bull Soc Zool Fr Evol Zool* 101(Suppl 4):15.
14. Cavallin M (1971) La "polyembryonie substitutive" et le probleme de l'origine de la lignée germinale chez le Phasme *Carausius morosus* Br. *C R Acad Sci Paris Ser III* 272: 462–465.
15. Strome S, Updike D (2015) Specifying and protecting germ cell fate. *Nat Rev Mol Cell Biol* 16(7):406–416.
16. Saitou M, Yamaji M (2012) Primordial germ cells in mice. *Cold Spring Harb Perspect Biol* 4(11):a008375.
17. Gunesdogan U, Magnusdottir E, Surani MA (2014) Primoridal germ cell specification: A context-dependent cellular differentiation event. *Philos Trans R Soc Lond, B* 369(1657):20130543.
18. Nakamura A, Seydoux G (2008) Less is more: Specification of the germline by transcriptional repression. *Development* 135(23):3817–3827.
19. Extavour CG (2007) Evolution of the bilaterian germ line: Lineage origin and modulation of specification mechanisms. *Integr Comp Biol* 47(5):770–785.
20. Evans T, Wade CM, Chapman FA, Johnson AD, Loose M (2014) Acquisition of germ plasm accelerates vertebrate evolution. *Science* 344(6180):200–203.
21. Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24(8):1586–1591.
22. Hurst LD (2002) The Ka/Ks ratio: Diagnosing the form of sequence evolution. *Trends Genet* 18(9):486.
23. Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125(1):1–15.
24. Lanfear R, Welch JJ, Bromham L (2010) Watching the clock: Studying variation in rates of molecular evolution between species. *Trends Ecol Evol* 25(9):495–503.
25. Whittle CA, Extavour CG (2016) Refuting the hypothesis that the acquisition of germ plasm accelerates animal evolution. *Nat Commun* 7:12637.
26. Crother BI, White ME, Johnson AD (2016) Diversification and germ-line determination revisited: Linking developmental mechanism with species richness. *Front Ecol Evol*, 10.3389/fevo.2016.00026.
27. Ricklefs RE (2007) Estimating diversification rates from phylogenetic information. *Trends Ecol Evol* 22(11):601–610.
28. Wertheim JO, Sanderson MJ (2011) Estimating diversification rates: How useful are divergence times? *Evolution* 65(2):309–320.
29. Alfaro ME, et al. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci USA* 106(32):13410–13414.
30. Swanson WJ, Vacquier VD (2002) The rapid evolution of reproductive proteins. *Nat Rev Genet* 3(2):137–144.
31. Singh R, Jagadeeshan S (2012) Sex and speciation: *Drosophila* reproductive tract proteins- twenty five years later. *Int J Evol Biol* 2012:191495.
32. Johnson AD, Alberio R (2015) Primordial germ cells: The first cell lineage or the last cells standing? *Development* 142(16):2730–2739.
33. Adkins RM, Gelke EL, Rowe D, Honeycutt RL (2001) Molecular phylogeny and divergence time estimates for major rodent groups: Evidence from multiple genes. *Mol Biol Evol* 18(5):777–791.
34. Jørgensen FG, et al. (2005) Comparative analysis of protein coding sequences from human, mouse and the domesticated pig. *BMC Biol* 3:2.
35. Toll-Riera M, Laurie S, Albà MM (2011) Lineage-specific variation in intensity of natural selection in mammals. *Mol Biol Evol* 28(1):383–398.
36. Fan X, Dougan ST (2007) The evolutionary origin of nodal-related genes in teleosts. *Dev Genes Evol* 217(11-12):807–813.
37. Innan H, Kondrashov F (2010) The evolution of gene duplications: Classifying and distinguishing between models. *Nat Rev Genet* 11(2):97–108.
38. Takahashi S, et al. (2006) Nodal-related gene Xnr5 is amplified in the *Xenopus* genome. *Genesis* 44(7):309–321.
39. Hellsten U, et al. (2007) Accelerated gene evolution and subfunctionalization in the pseudotetraploid frog *Xenopus laevis*. *BMC Biol* 5:31.
40. Crow KD, Smith CD, Cheng JF, Wagner GP, Amemiya CT (2012) An independent genome duplication inferred from Hox paralogs in the American paddlefish—a representative basal ray-finned fish and important comparative reference. *Genome Biol Evol* 4(9):937–953.
41. Glasauer SM, Neuhauss SC (2014) Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol Genet Genomics* 289(6):1045–1060.

42. Smith JJ, et al. (2009) Genic regions of a large salamander genome contain long introns and novel genes. *BMC Genomics* 10:19.
43. Otto SP, Hastings IM (1998) Mutation and selection within the individual. *Genetica* 102-103(1-6):507–524.
44. Extavour C, García-Bellido A (2001) Germ cell selection in genetic mosaics in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 98(20):11341–11346.
45. Choi SK, Yoon SR, Calabrese P, Arnheim N (2008) A germ-line-selective advantage rather than an increased mutation rate can explain some unexpectedly common human disease mutations. *Proc Natl Acad Sci USA* 105(29):10143–10148.
46. Yoon SR, et al. (2013) Age-dependent germline mosaicism of the most common noonan syndrome mutation shows the signature of germline selection. *Am J Hum Genet* 92(6):917–926.
47. Ellegren H (2007) Characteristics, causes and evolutionary consequences of male-biased mutation. *Proc Biol Sci* 274(1606):1–10.
48. Ellegren H, Parsch J (2007) The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* 8(9):689–698.
49. Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proc R Soc Lond B Biol Sci* 205(1161):581–598.
50. Manhart M, Morozov AV (2015) Protein folding and binding can emerge as evolutionary spandrels through structural coupling. *Proc Natl Acad Sci USA* 112(6): 1797–1802.
51. Gilbert SF (2013) *Developmental Biology* (Sinauer Associates, Sunderland, MA), 10th Ed, p 750.
52. Bashirullah A, Cooperstock RL, Lipshitz HD (1998) RNA localization in development. *Annu Rev Biochem* 67:335–394.
53. Schier AF, Talbot WS (2005) Molecular genetics of axis formation in zebrafish. *Annu Rev Genet* 39:561–613.
54. Tao Q, et al. (2005) Maternal wnt11 activates the canonical wnt signaling pathway required for axis formation in *Xenopus* embryos. *Cell* 120(6):857–871.
55. Jedrusik A, et al. (2008) Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. *Genes Dev* 22(19):2692–2706.
56. Skamagki M, Wicher KB, Jedrusik A, Ganguly S, Zernicka-Goetz M (2013) Asymmetric localization of Cdx2 mRNA during the first cell-fate decision in early mouse development. *Cell Reports* 3(2):442–457.
57. Jedrusik A, et al. (2010) Maternally and zygotically provided Cdx2 have novel and critical roles for early development of the mouse embryo. *Dev Biol* 344(1):66–78.
58. Jedrusik A, Cox A, Wicher KB, Glover DM, Zernicka-Goetz M (2015) Maternal-zygotic knockout reveals a critical role of Cdx2 in the morula to blastocyst transition. *Dev Biol* 398(2):147–152.
59. Blij S, Frum T, Akyol A, Fearon E, Ralston A (2012) Maternal Cdx2 is dispensable for mouse development. *Development* 139(21):3969–3972.
60. Dietrich JE, Hiiragi T (2007) Stochastic patterning in the mouse pre-implantation embryo. *Development* 134(23):4219–4231.
61. Frankenberg S, Zernicka-Goetz M (2004) Breaking radial symmetry—amniote type. *Gastrulation*, ed Stern C (Cold Spring Harbor Lab Press, New York).
62. Arnold SJ, Robertson EJ (2009) Making a commitment: Cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol* 10(2):91–103.
63. Maître JL, et al. (2016) Asymmetric division of contractile domains couples cell positioning and fate specification. *Nature* 536(7616):344–348.
64. Momose T, Houliston E (2007) Two oppositely localised frizzled RNAs as axis determinants in a cnidarian embryo. *PLoS Biol* 5(4):e70.
65. Haillot E, Molina MD, Lapraz F, Lepage T (2015) The maternal Maverick/GDF15-like TGF-β ligand panda directs dorsal-ventral axis formation by restricting nodal rxpression in the sea urchin embryo. *PLoS Biol* 13(9):e1002247.
66. Lynch JA, et al. (2011) The phylogenetic origin of *oskar* coincided with the origin of maternally provisioned germ plasm and pole cells at the base of the Holometabola. *PLoS Genet* 7(4):e1002029.
67. Goldberg EE, Igić B (2008) On phylogenetic tests of irreversible evolution. *Evolution* 62(11):2727–2741.
68. Bull JJ, Charnov EL (1985) On irreversible evolution. *Evolution* 39:1149–1155.
69. Barrett SC (2013) The evolution of plant reproductive systems: How often are transitions irreversible? *Proc Biol Sci* 280(1765):20130913.
70. Smith SD, Rausher MD (2011) Gene loss and parallel evolution contribute to species difference in flower color. *Mol Biol Evol* 28(10):2799–2810.
71. Nakamura T, Extavour CG (2016) The transcriptional repressor Blimp-1 acts downstream of BMP signaling to generate primordial germ cells in the cricket *Gryllus bimaculatus*. *Development* 143(2):255–263.
72. Deshpande G, et al. (2014) BMP signaling and the maintenance of primordial germ cell identity in *Drosophila* embryos. *PLoS One* 9(2):e88847.
73. Saffman EE, Lasko P (1999) Germline development in vertebrates and invertebrates. *Cell Mol Life Sci* 55(8-9):1141–1163.
74. Takamura K, Fujimura M, Yamaguchi Y (2002) Primordial germ cells originate from the endodermal strand cells in the ascidian *Ciona intestinalis*. *Dev Genes Evol* 212(1): 11–18.
75. Stern DL (2013) The genetic causes of convergent evolution. *Nat Rev Genet* 14(11): 751–764.
76. Harris WA (1997) Pax-6: Where to be conserved is not conservative. *Proc Natl Acad Sci USA* 94(6):2098–2100.
77. Futuyma DJ (1997) *Evolutionary Biology* (Sinauer, Sunderland, MA), 3rd Ed.
78. Hillis DM, Heller C, Sadava DE (2010) Principles of Life (Sinauer and W. H. Freeman, New York), p 915.
79. Mahowald AP (2001) Assembly of the *Drosophila* germ plasm. *Int Rev Cytol* 203: 187–213.
80. Smith JL, Wilson JE, Macdonald PM (1992) Overexpression of *oskar* directs ectopic activation of *nanos* and presumptive pole cell formation in *Drosophila* embryos. *Cell* 70(5):849–859.
81. Ephrussi A, Lehmann R (1992) Induction of germ cell formation by *oskar*. *Nature* 358(6385):387–392.
82. Ewen-Campen B, Srouji JR, Schwager EE, Extavour CG (2012) *oskar* predates the evolution of germ plasm in insects. *Curr Biol* 22(23):2278–2283.
83. Bontems F, et al. (2009) Bucky ball organizes germ plasm assembly in zebrafish. *Curr Biol* 19(5):414–422.
84. Heim AE, et al. (2014) Oocyte polarity requires a Bucky ball-dependent feedback amplification loop. *Development* 141(4):842–854.
85. Marlow FL, Mullins MC (2008) Bucky ball functions in Balbiani body assembly and animal-vegetal polarity in the oocyte and follicle cell layer in zebrafish. *Dev Biol* 321(1):40–50.
86. Leacock SW, Reinke V (2008) MEG-1 and MEG-2 are embryo-specific P-granule components required for germline development in *Caenorhabditis elegans*. *Genetics* 178(1):295–306.
87. Wang JT, et al. (2014) Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *eLife* 3:e04591.
88. Chang CW, et al. (2011) Anterior-posterior axis specification in *Drosophila* oocytes: Identification of novel *bicoid* and *oskar* mRNA localization factors. *Genetics* 188(4): 883–896.
89. Updike D, Strome S (2010) P granule assembly and function in *Caenorhabditis elegans* germ cells. *J Androl* 31(1):53–60.
90. Langdon YG, Mullins MC (2011) Maternal and zygotic control of zebrafish dorsoventral axial patterning. *Annu Rev Genet* 45:357–377.
91. Elena SF, Cooper VS, Lenski RE (1996) Punctuated evolution caused by selection of rare beneficial mutations. *Science* 272(5269):1802–1804.
92. Marlow F (2015) Primordial germ cell specification and migration. *F1000 Res* 4:1462.
93. Wourms JP, Lombardi J (1992) Reflections on the evolution of piscine viviparity. *Integr Comp Biol* 32(2):276–293.
94. Muller HJ (1932) Further studies on the nature and causes of gene mutations. *Sixth Int Congress Genet* 1:213–255.
95. Bromham L (2009) Why do species vary in their rate of molecular evolution? *Biol Lett* 5(3):401–404.
96. Rodgers K, McVey M (2016) Error-prone repair of DNA double-strand breaks. *J Cell Physiol* 231(1):15–24.
97. Datta A, Jinks-Robertson S (1995) Association of increased spontaneous mutation rates with high levels of transcription in yeast. *Science* 268(5217):1616–1619.
98. Kondrashov FA, Kondrashov AS (2010) Measurements of spontaneous rates of mutations in the recent past and the near future. *Philos Trans R Soc Lond B Biol Sci* 365(1544):1169–1176.
99. Kondrashov AS (1998) Measuring spontaneous deleterious mutation process. *Genetica* 102-103(1-6):183–197.
100. Drake JW, Charlesworth B, Charlesworth D, Crow JF (1998) Rates of spontaneous mutation. *Genetics* 148(4):1667–1686.
101. Lynch M (2010) Evolution of the mutation rate. *Trends Genet* 26(8):345–352.
102. Sung W, et al. (2016) Evolution of the insertion-deletion mutation rate across the tree of life. *G3 (Bethesda)* 6(8):2583–2591.
103. Haldane JBS (1947) The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. *Ann Eugen* 13(4):262–271.
104. Campbell CD, Eichler EE (2013) Properties and rates of germline mutations in humans. *Trends Genet* 29(10):575–584.
105. Rahbari R, et al.; UK10K Consortium (2016) Timing, rates and spectra of human germline mutation. *Nat Genet* 48(2):126–133.
106. Kong A, et al. (2012) Rate of de novo mutations and the importance of father's age to disease risk. *Nature* 488(7412):471–475.
107. Drost JB, Lee WR (1995) Biological basis of germline mutation: Comparisons of spontaneous germline mutation rates among drosophila, mouse, and human. *Environ Mol Mutagen* 25(Suppl 26):48–64.
108. Lynch M (2010) Rate, molecular spectrum, and consequences of human mutation. *Proc Natl Acad Sci USA* 107(3):961–968.
109. Keightley PD, et al. (2009) Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Res* 19(7): 1195–1201.
110. Smeds L, Qvarnström A, Ellegren H (2016) Direct estimate of the rate of germline mutation in a bird. *Genome Res* 26(9):1211–1218.
111. Gillooly JF, McCoy MW, Allen AP (2007) Effects of metabolic rate on protein evolution. *Biol Lett* 3(6):655–659.
112. Ségurel L, Wyman MJ, Przeworski M (2014) Determinants of mutation rate variation in the human germline. *Annu Rev Genomics Hum Genet* 15:47–70.
113. Mohlhenrich ER, Mueller RL (2016) Genetic drift and mutational hazard in the evolution of salamander genomic gigantism. *Evolution* 70(12):2865–2878.
114. Dunn CW, et al. (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452(7188):745–749.

Supporting Information (SI): Appendix

Causes and evolutionary consequences of primordial germ cell specification mode in metazoans

CA Whittle[a] and CG Extavour[a,b]

[a] Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge MA 02138, USA

[b] Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge MA 02138, USA

Table S1. Summary of the hypotheses and proposals pertaining to PGC specification mode in metazoans.

	Supporting Evidence	Refuting Evidence
Hypotheses pertaining to causes of PGC specification mode in metazoans		
●***PGC specification hypothesis*** (1, 2): Inheritance mode evolves convergently due to the liberation of selective constraint on the soma and somatic gene regulatory networks (SGRNs), leading to enhanced "evolvability", accelerating an organism's gene sequence evolution, and enhancing speciation. Induction is linked to constrained, or slowed, evolution and reduced speciation.	Faster rates of nonsynonymous substitution under germ plasm in non-phylogenetically independent contrasts of vertebrates claimed as support (1, 2); the limitations of those conclusions are cited in the main text. High species richness in some clades with germ plasm, as compared to those with induction (1, 3). Developmental innovations, and variation in gene pathways (such as *Nodal* in Xenopus) under inheritance mode (1).	Support of the null hypothesis (no effect) of inheritance on selective constraint and protein sequence evolution based on nonsynonymous to synonymous substitution (dN/dS) rates in phylogenetically independent animal genera (4). Species richness might not reflect speciation rates; the latter yields cases showing no effect of germ plasm (5). Examples of germ plasm clades that are species poor relative to induction clades (4).
● ***Deterministic-stochastic PGC hypothesis*** (6): PGC specification mode is classified as determinative ("early" PGC specification; effectively those with germ plasm and induction taxa with early PGC formation, namely rodents) or stochastic ("late" specification, presumably most induction taxa). Particularly focused on rodents, including mice, which have "early" specification under induction. The hypothesis posits the determinative mode is linked to liberation of selective constraint on SGRNs, leading to enhanced "evolvability", fast	Higher species richness in deterministic clades, including an "early" inductive mammalian clade (rodents), as compared to those with stochastic mode (6). Developmental innovations (such as egg-cylinder in mouse), and variation in gene pathways (such as *Nodal* in Xenopus) under determinative mode (6). Mice exhibiting relatively fast sequence divergence among mammals for some genes taken as support (6, 7).	Insufficient empirical study to date. Additional study of induction taxa with "early" PGC specification required beyond rodents (if identifiable). Examples in animals suggest taxa with germ plasm ("deterministic" under this hypothesis) do not consistently exhibit accelerated sequence evolution, or liberated constraints (4). Mice do not exhibit elevated dN/dS values, as compared to other mammals such as humans and pigs (8), inconsistent with release of constraint. See also refuting evidence for species richness under the PGC specification hypothesis above.

evolution, and enhanced speciation. Stochastic mode is linked to constrained, or slowed evolution, and reduced speciation.		
● ***Germ plasm is a side-effect, or spandrel***: Germ plasm has evolved as a side-effect, or spandrel (9), of a general shift towards the usage of maternal determinants for body patterning.	Speculative hypothesis: consistent with co-occurrence of germ plasm and maternal determinants used in models such as flies, zebrafish and frogs (10-12) and the use of signalling and cell-cell interactions for both PGC specification and embryonic patterning in mammals (13-16).	May be an artefact of model taxa studied to date. Some models of cnidarians and echinoderms might use inductive PGC formation and maternal determinants for patterning (17, 18).
● ***PGC specification mode influences the germ line mutation rate***. If the mutation rate is lower under inheritance mode, it could be advantageous and cause convergent evolution of germ plasm. Early establishment of PGCs (low cell division and transcription), reduced propensity for mutation in the pre-PGC stages, and/or greater cell-lineage selection in the germ line under inheritance (than induction) mode might favor a lowered mutation rate per generation.	Speculative hypothesis: Consistent with data that the pre-PGC stage can influence the rate of mutation in animals, a phase when inheritance and induction modes differ in development and genetics (19, 20). Concordant with some (very limited) mutation rate data available to date from animal/insect models with inheritance and induction (see for example 21, 22, 23).	Insufficient empirical data to date to test rigorously.
Proposals related to the evolution of PGC specification modes		
● Inheritance has convergently evolved from induction; thus inheritance is the derived state	This scenario appears probable given inheritance typically evolves from induction across the metazoan phylogeny, and that	Convergent evolution from inheritance to induction mode across metazoans cannot fully excluded. Putative examples of transitions from

and induction is the ancestral state.	induction is characteristic of the basally branching metazoans (24).	inheritance to induction in some insect orders, although these may be a reversion to an ancestral induction state in a last common insect ancestor (25, 26). In addition, germ plasm may evolve as a side effect, rather than convergent, or adaptive, evolution (see above).
● The transition to inheritance mode is Irreversible: “Dollo’s Law.”	Transitions from induction to inheritance mode are typical across the metazoan phylogeny (24). Absence/loss of induction mode capabilities in species with germ plasm (27-29). Evidence of reverse transitions from inheritance to induction mode is rare or possibly absent in the metazoan phylogeny (24, but see 30, 31).	Plausible cases of taxa with both inheritance and induction modes (e.g. 30, 31) (Note: this might also be indicative of a transition period to from induction to inheritance.). Possibility of transitions from inheritance to induction mode, (or reversion to inheritance from induction), in some insect orders (26).
● Inheritance mode arises via distinct genetic and developmental mechanisms.	As examples, germ plasm assembly is controlled by *oskar* in fruit flies (32-34), *bucky ball* in zebrafish (35-37) and *MEG/PGL* in roundworms (38, 39). Biological properties of germ plasm differ among inheritance taxa (1, 40).	None.
● Reproductive lifestyles, including oviparity or viviparity, are linked to the transition from induction to inheritance mode. For instance, a transition to germ plasm under stress may be particularly favored under oviparity.	Speculative: An enhanced propensity for a transition to germ plasm under oviparity agrees with some available data from animals (see main text for discussion and citations)[a].	Various taxa have both viviparity and germ plasm. However, these may have evolved germ plasm under oviparity (before switch to viviparity)[a]. Also, germ plasm could have some advantages under both oviparity and viviparity. Future testing is warranted.

● The germ line-soma divide favors a lower mutation rate in the germ line (23, 41, 42). This phenomenon occurs regardless of PGC specification mode.	Lowered mutation rates in germ line than soma in animal models such as humans, mice, flies, fish (22, 23), spanning both induction and inheritance taxa[a]. May result from phases of mitotic and transcriptional quiescence in germ line development (42, 43).	Further study in broader range of taxa warranted.
● PGC Specification Mode Shapes Molecular Evolution	Speculative: Perhaps may occur on a much smaller scale than predicted by the PGC specification hypothesis (1, 2). Potentially may influence genes involved in PGC specification; particularly genes related to germ plasm (if adaptive) may evolve rapidly.	Insufficient empirical study to date.

[a] For detailed arguments and citations see main text.

Supporting Text

Section 1: Considering the Possibility of Inheritance as the Ancestral Mode

Given the distribution of inheritance and induction PGC specification modes across animals, the preponderance of evidence suggests that the latter comprises the ancestral state (24). Nonetheless, it is worth considering the possibility that inheritance comprises the ancestral state, has been lost from the most basally branching extant lineages (sponges, ctenophores, and many cnidarians) and from most extant phyla, and has undergone many transitions to induction over animal evolutionary history. One could argue, for instance, that the generation of germ plasm before successful fertilization is a costly and inefficient process, as if the egg remains unfertilized this comprises wasted biochemical resources. In this regard, a switch from inheritance to induction could be favored as it limits PGC specification costs to after fertilization and embryo formation (from signalling within the embryo). Additional evidence could also be considered consistent with inheritance as the ancestral model. For instance, the BMP pathway, which is crucial for PGC formation in mice (44), has not been found to be involved in inductive PGC specification in most other induction animals, with the exception of crickets and salamanders (45-48), and perhaps humans and pigs based on stem cell differentiation studies (6, 49, 50). (We note, however, that to our knowledge, few or no studies have been undertaken outside of these taxa to explicitly determine whether or not BMP signalling, or any other signalling pathway, induces or is required for germ cell formation.). Thus, whilst the general functional properties of induction are mostly conserved across metazoans (24), it is still unclear to what extent the specific molecular regulators in the induction pathway are well conserved. In this respect, the pattern of induction could be consistent with convergent functional evolution from germ plasm, particularly if different genetic pathways are found to be operative in inducing PGCs in different induction lineages (51). Further molecular data on the upstream regulatory

genes involved in the induction pathway in a greater range of species will be needed to fully ascertain the degree of conservation across induction regulators, including BMP signaling, in metazoans.

The strongest evidence to date supporting an ancestral state of induction, and convergent evolution of germ plasm is that, with very few possible exceptions (26), the majority of available data to date points towards inheritance being far more widespread and frequent than induction across both protostomes and deuterostomes, and that the basally branching animal lineages likely use the inheritance mode almost exclusively (24, but see 52). Thus, convergent evolution of germ plasm from induction currently comprises the most plausible scenario underlying the distribution of PGC specification modes across metazoans.

Section 2: Limitations of a Study on the PGC-Specification Hypothesis

We note the limitations of the Evans et al. (2014) (2) investigation in detail here. In the main assessment of four-species trees (e.g. anuran, urodele, mammal and outgroup), which was the primary method chosen to examine the patterns of protein coding sequence evolution in inheritance versus induction species, results from only two of the four pairs of vertebrates under study (see above) supported the hypothesis. Further, the findings were based on examination of nonsynonymous coding-DNA substitutions (dN) among taxa, and excluded the synonymous site changes (dS; excluded due to saturation) as well as the ratio dN/dS, which is needed to assess variation (or "liberation") in selection pressures (53, 54). Without dS, one cannot ascertain whether differences in dN among lineages results from mutational or selection pressures (53, 54). Importantly, the main approach of four-species trees and the follow-up comparisons using relative rate tests (which are paired contrasts of evolutionary rates between two related paired taxa using an outgroup) involved many paired contrasts of inheritance and induction taxa that were not phylogenetically independent, but were nonetheless treated as independent data points (as an example, many species of anurans were compared to species of urodeles), which is pseudoreplication (55, 56). This approach is strongly cautioned against, as it can give the false impression of robust general relationships due to the repeated non-independent sampling of taxa (or, phylogenetic branches) with shared ancestry (55, 56). Small gene sample sizes for some taxa, biases towards using genes involved in specific tissue-types and/or functions, and the exclusion of invertebrates, which comprise the vast majority of animal life, were additional features of the Evans et al. approach that might have contributed to conclusions drawn in that study (4).

Section 3: PGC-Specification may not be Linked to Speciation

Here, we provide examples suggesting that germ plasm might not be a causative factor linked to high speciation rates in some animal groups. For instance, an assessment of 44 clades of jawed vertebrates using diversification rate analysis that includes birth-death rates, clade-age and DNA sequence analysis, suggests that despite having reported species numbers that differ by an order of magnitude, Anura (frogs; inheritance) exhibit divergence rates that are not only typical for vertebrates, but also similar to, rather than higher than, those of its sister taxon Caudata (salamanders; induction) (5). Based on species richness alone, it had been proposed that these two clades differ markedly in their level of speciation, with higher diversification in the inheritance taxon (1, 3), and thus species numbers were taken as support for the PGC specification hypothesis. However, at least in this particular case, the diversification-rate and species richness findings do not concur, and the former approach does not support higher speciation under inheritance. Additional diversification rate studies using those clades with reported high species richness under germ plasm (1, 3) would be informative in ascertaining whether an effect remains to be observed in the same lineages under a more exacting assessment of speciation. Furthermore, diversification data from jawed vertebrates showed that taxa exhibiting both inheritance (e.g. birds, teleosts) and induction (e.g. lizards, eutherian mammals) exhibit accelerated (non-typical) diversification rates (5) in this group, and thus speciation rates appear unconnected to PGC specification mode in these vertebrates. Moreover, species richness itself appears unconnected to PGC specification mode in multiple invertebrate lineages, also suggesting an inconsistency with the PGC specification hypothesis in those animal groups (4).

Reproductive Isolation

As noted in the main text, an important aspect of the prediction of enhanced speciation under germ plasm (1, 2) that warrants contemplation is the mechanism of reproductive isolation. For example, the notion that greater speciation occurs under germ plasm would require a proposed mechanism for enhanced instances of reproductive isolation, which underlie speciation episodes. It can be speculated here that a suitable argument one could employ might be that germ plasm itself comprises a rapidly diverging trait, causing egg properties and/or genes linked to fertilization (c.f. 57) to diverge from sperm in a fast manner, and causing incompatibility. In turn, another speculative, yet conceivable, mechanism consistent with the hypothesis might involve the rapid divergence of the male and female somatic reproductive organs under germ plasm (and underlying sex and reproduction-related genes (SSR), 58), impeding mating or post-mating reproductive success, and thus causing reproductive isolation. Taken in combination with factors such as geographic isolation and sexual selection that are also believed to often be involved in speciation (59, 60), these could comprise conceivable pathways to enhanced speciation under germ plasm.

Section 4: Molecular Evolution and PGC Specification

Whilst counter to the PGC specification hypothesis (1, 2), evidence suggests germ plasm might not be linked to a broad propensity for rapid evolution of protein sequences (4), it could nonetheless be connected to the evolution of genes that are specifically involved in PGC specification mechanisms. In Drosophila (inheritance), for instance, this should include the gene *oskar*. If *oskar* is indeed the result of an adaptive mutation needed for or facilitating the transition to inheritance mode, then one would anticipate it to exhibit rapid evolution, as evidenced by high nonsynonymous to synonymous (dN/dS) substitution rates (53). Available data does suggest that *oskar* evolves rapidly within the Drosophila genus. First, the *oskar* ortholog of *D. virilis* fails to rescue *oskar* loss of function mutants in *D. melanogaster* (61), suggesting rapid functional divergence in this genus. Further, dN/dS indicates that sites in at least two domains of *oskar* (LASP binding and Long Osk domain) have experienced positive selection (62), consistent with adaptive evolution. In contrast to inheritance, if induction is indeed the ancestral state and well conserved, the genes regulating this PGC specification mode, such as BMP proteins in mouse (63), might be expected to evolve under high purifying selection. Evidence suggests that BMP2 has been subjected to bursts of positive selection (at specific sites) in birds (inheritance), less apparent in some mammals such as mouse (induction) (64), consistent with a possible reduced propensity for positive selection on this gene under induction. However, as BMPs play pleiotropic roles in the development of all animals, it is not possible to isolate the precise effect of induction that could cause greater conservation (64-66). Together, the limited data from *oskar* and BMP genes suggest that it will be worthwhile to assess the molecular evolution of additional genes specifically expressed, and likely functionally involved (63), in pre-PGC formation cells and early-post-PGC stages. In this manner, it may be revealed whether

inheritance and induction exert selection pressures in early germ line establishment, which in turn could shape the molecular evolution of genes involved in PGC specification.

Section 5: Reproductive Lifestyles and PGC-Specification Mode

We discuss here the putative relationship between PGC specification and reproductive lifestyles, particularly viviparity and oviparity, and whether inheritance may be more apt to evolve under oviparity. As an example, in typically oviparous amphibian (class Amphibia) organisms, including frogs (inheritance) and salamanders (induction) (mostly oviparous (67)), a temporary shift to stressful conditions impairing reproductive output early in the frog lineage (in the frog common ancestor) but not occurring in salamander ancestors, could promote an irreversible switch from induction to germ plasm, if germ plasm enhanced germ line and embryo survival. Numerous cases of a switch from induction to germ plasm evolution have occurred in typically oviparous organisms (68), including birds, arthropods (particularly insects), and teleost fish (1, 24, 26). In contrast, transitions from induction to germ plasm appear less evident in viviparous taxa: for instance, most mammals (99% of species are viviparous (68)), lizards (approximately 80% of species are viviparous (69, 70)), and the Cnidaria, a basally branching metazoan taxon (some viviparous (71)), nearly all exhibit induction in taxa studied to date (2, 24, 26, 52). It should be noted that teleosts, which use germ plasm (1) and are typically oviparous, include some species that are viviparous (72). This however, likely reflects a transition to viviparity under germ plasm (rather than a transition to germ plasm under viviparity), as teleosts as a group are believed to specify PGCs using inheritance (and thus inheritance is likely an ancestral state to the clade; 1). Similarly, germ plasm has been reported in snakes, nematodes and some Ascidians (24) wherein the families/genera are typically oviparous, but some exhibit viviparity (69-71, 73, 74). In these cases, as many species (per clade) are not viviparous, it is plausible that germ plasm evolved before within-clade transitions to viviparity from the ancestral oviparous state (which are common transitions in certain animal clades (70)), suggesting the

switch to germ plasm occurred under oviparity. Confirmation of the ancestral state per clade as oviparous before the transition to germ plasm would be needed for verification that the switch to germ plasm occurred under oviparity; nonetheless, an oviparous ancestry is implied by the fact that oviparity is considered the ancestral primitive state in animals, and viviparity the advanced and derived state within the clades where it is observed (75, 76). Collectively, these anecdotes might suggest that a switch to germ plasm is more commonly beneficial to embryo survival and organismal fitness under stress in oviparous (more than in viviparous) taxa. As the number of animal taxa wherein PGC specification mechanisms have been studied and identified at the molecular mechanistic level remains limited (24), further investigation will be needed to ascertain how commonly the evolution of germ plasm correlates with oviparity and viviparity, and whether germ plasm, may be particularly advantageous under one type of reproduction mode.

Section 6: Separation of the Germ Line and Soma

Under a model wherein many mutations in coding DNA are deleterious (22, 77), and the germ line gives rise to all genetically heritable mutations, one may predict that for germ line segregation to be advantageous the germ lines should exhibit a lower mutation rate per cell division than the soma (41, 42). Whilst comparative data on mutation rates in the germ lines and soma remain sparse, available findings are suggestive of such an effect. For example, estimates in humans using B and T lymphocytes, fibroblasts, retina, and intestinal cells have shown a 4- to 25-fold higher rate of mutation per base-pair per cell division in somatic cells than in the germ lines (22, 23), consistent with a selective reduction in the mutation rate in the germ cell lineage. Further, at early sexual maturity in humans (15 years), the soma has been found to have between a 10- and 100-fold higher number of accumulated coding DNA mutations than germ line cells (22). Similarly, in mice, the soma has been reported to have an elevated mutation rate as compared to germ cells, with 2- to 10-fold higher (or even higher) mutation rate per site per cell division than the male germ line, and a 2 to 6-fold higher level of accumulated mutations at maturity (22, 23, 41, 78). In terms of absolute time scale, in *D. melanogaster* somatic mutation rates have been found to be about 80-fold higher than rates in the germ line, and medaka fish also exhibit higher germ line than somatic mutation rates (reviewed in 23). Collectively, the findings from these animals are consistent with a selective advantage of lower mutation rates in the germ line than in the soma, thus favoring the evolution of the germ line-soma divide. It can be proposed that the lowered germ line mutation rate might partly result from the distinct cellular dynamics of PGCs and germ cells, which could include the typically low (or absent) mitotic activity that occurs between PGC specification and gametogenesis (reviewed in 42) limiting replication errors, and/or from transcriptional quiescence or limited transcriptional activity in

PGCs and germ cells (42, 43), all of which would limit the frequency of transcription-mediated mutations (79, 80).

Supplementary References

1. Johnson AD, Richardson E, Bachvarova RF, & Crother BI (2011) Evolution of the germ line-soma relationship in vertebrate embryos. *Reproduction* 141(3):291-300.
2. Evans T, Wade CM, Chapman FA, Johnson AD, & Loose M (2014) Acquisition of germ plasm accelerates vertebrate evolution. *Science* 344(6180):200-203.
3. Crother BI, White ME, & Johnson AD (2016) Diversification and Germ-Line Determination Revisited: Linking Developmental Mechanism with Species Richness. *Frontiers in Ecology and Evolution* 2016.
4. Whittle CA & Extavour CG (2016) Refuting the hypothesis that the acquisition of germ plasm accelerates animal evolution. *Nat. Comm.* 7:12637.
5. Alfaro ME*, et al.* (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc. Natl. Acad. Sci. USA* 106(32):13410-13414.
6. Johnson AD & Alberio R (2015) Primordial germ cells: the first cell lineage or the last cells standing? *Development* 142(16):2730-2739.
7. Adkins RM, Gelke EL, Rowe D, & Honeycutt RL (2001) Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Mol. Biol. Evol.* 18(5):777-791.
8. Jorgensen FG*, et al.* (2005) Comparative analysis of protein coding sequences from human, mouse and the domesticated pig. *BMC Biol* 3:2.
9. Gould SJ & Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B, Biological sciences* 205(1161):581-598.
10. Bashirullah A, Cooperstock RL, & Lipshitz HD (1998) RNA localization in development. *Annu. Rev. Biochem.* 67:335-394.
11. Schier AF & Talbot WS (2005) Molecular genetics of axis formation in zebrafish. *Annu. Rev. Genet.* 39:561-613.
12. Tao Q*, et al.* (2005) Maternal wnt11 activates the canonical wnt signaling pathway required for axis formation in *Xenopus* embryos. *Cell* 120(6):857-871.
13. Dietrich JE & Hiiragi T (2007) Stochastic patterning in the mouse pre-implantation embryo. *Development* 134(23):4219-4231.
14. Maitre JL*, et al.* (2016) Asymmetric division of contractile domains couples cell positioning and fate specification. *Nature* 536(7616):344-348.
15. Frankenberg S & Zernicka-Goetz M (2004) Breaking Radial Symmetry - Amniote Type. *Gastrulation*, ed Stern C (Cold Spring Harbor Laboratory Press, New York).
16. Arnold SJ & Robertson EJ (2009) Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol* 10(2):91-103.
17. Momose T & Houliston E (2007) Two oppositely localised frizzled RNAs as axis determinants in a cnidarian embryo. *PLoS Biol.* 5(4):e70.
18. Haillot E, Molina MD, Lapraz F, & Lepage T (2015) The Maternal Maverick/GDF15-like TGF-beta Ligand Panda Directs Dorsal-Ventral Axis Formation by Restricting Nodal Expression in the Sea Urchin Embryo. *PLoS Biol.* 13(9):e1002247.
19. Drost JB & Lee WR (1995) Biological basis of germline mutation: comparisons of spontaneous germline mutation rates among drosophila, mouse, and human. *Environ. Mol. Mutagen.* 25 Suppl 26:48-64.

20. Rahbari R*, et al.* (2016) Timing, rates and spectra of human germline mutation. *Nat. Genet.* 48(2):126-133.
21. Keightley PD*, et al.* (2009) Analysis of the genome sequences of three Drosophila melanogaster spontaneous mutation accumulation lines. *Genome Res.* 19(7):1195-1201.
22. Lynch M (2010) Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl. Acad. Sci. USA* 107(3):961-968.
23. Lynch M (2010) Evolution of the mutation rate. *Trends Genet.* 26(8):345-352.
24. Extavour CG & Akam ME (2003) Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* 130(24):5869-5884.
25. Ewen-Campen B, Donoughe S, Clarke DN, & Extavour CG (2013) Germ cell specification requires zygotic mechanisms rather than germ plasm in a basally branching insect. *Curr. Biol.* 23(10):835-842.
26. Lynch JA*, et al.* (2011) The Phylogenetic Origin of *oskar* Coincided with the Origin of Maternally Provisioned Germ Plasm and Pole Cells at the Base of the Holometabola. *PLoS Genetics* 7(4):e1002029.
27. Saffman EE & Lasko P (1999) Germline development in vertebrates and invertebrates. *Cell. Mol. Life Sci.* 55(8-9):1141-1163.
28. Eddy EM (1984) Origin of the germ cell line. *Ultrastructure of Reproduction: Gametogenesis, Fertilization and Embryogenesis*, eds Blerkom Jv & Motta PM (Springer USA, Boston, MA).
29. Hashimoto Y*, et al.* (2004) Localized maternal factors are required for zebrafish germ cell formation. *Dev. Biol.* 268(1):152-161.
30. Modrell M*, et al.* (in preparation) Germline replacement following ablation of the primordial germ cells in *Parhyale hawaiensis*.
31. Takamura K, Fujimura M, & Yamaguchi Y (2002) Primordial germ cells originate from the endodermal strand cells in the ascidian *Ciona intestinalis*. *Dev. Genes Evol.* 212(1):11-18.
32. Ephrussi A & Lehmann R (1992) Induction of germ cell formation by *oskar*. *Nature* 358(6385):387-392.
33. Smith JL, Wilson JE, & Macdonald PM (1992) Overexpression of *oskar* directs ectopic activation of *nanos* and presumptive pole cell formation in *Drosophila* embryos. *Cell* 70(5):849-859.
34. Mahowald AP (2001) Assembly of the *Drosophila* germ plasm. *International review of cytology* 203:187-213.
35. Marlow FL & Mullins MC (2008) Bucky ball functions in Balbiani body assembly and animal-vegetal polarity in the oocyte and follicle cell layer in zebrafish. *Dev. Biol.* 321(1):40-50.
36. Bontems F*, et al.* (2009) Bucky ball organizes germ plasm assembly in zebrafish. *Curr. Biol.* 19(5):414-422.
37. Heim AE*, et al.* (2014) Oocyte polarity requires a Bucky ball-dependent feedback amplification loop. *Development* 141(4):842-854.
38. Leacock SW & Reinke V (2008) MEG-1 and MEG-2 are embryo-specific P-granule components required for germline development in Caenorhabditis elegans. *Genetics* 178(1):295-306.
39. Wang JT*, et al.* (2014) Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans. *eLife* 3:e04591.

40. Marlow F (2015) Primordial Germ Cell Specification and Migration. *F1000Res* 4.
41. Walter CA, Intano GW, McCarrey JR, McMahan CA, & Walter RB (1998) Mutation frequency declines during spermatogenesis in young mice but increases in old mice. *Proc. Natl. Acad. Sci. USA* 95(17):10015-10019.
42. Extavour CG (2007) Evolution of the bilaterian germ line: lineage origin and modulation of specification mechanisms. *Integr. Comp. Biol.* 47(5):770-785.
43. Strome S & Updike D (2015) Specifying and protecting germ cell fate. *Nat. Rev. Mol. Cell. Biol.* 16(7):406-416.
44. Saitou M & Yamaji M (2012) Primordial germ cells in mice. *Cold Spring Harb Perspect Biol* 4(11):a008375.
45. Nakamura T & Extavour CG (2016) The transcriptional repressor Blimp-1 acts downstream of BMP signaling to generate primordial germ cells in the cricket *Gryllus bimaculatus*. *Development* 143(2):255-263.
46. Johnson AD*, et al.* (2003) Evolution of predetermined germ cells in vertebrate embryos: implications for macroevolution. *Evol. Dev.* 5(4):414-431.
47. Chatfield J*, et al.* (2014) Stochastic specification of primordial germ cells from mesoderm precursors in axolotl embryos. *Development* 141(12):2429-2440.
48. Donoughe S*, et al.* (2014) BMP signaling is required for the generation of primordial germ cells in an insect. *Proc. Natl. Acad. Sci. USA* 111(11):4133-4138.
49. Kee K, Gonsalves JM, Clark AT, & Pera RAR (2006) Bone morphogenetic proteins induce germ cell differentiation from human embryonic stem cells. *Stem Cells Dev* 15(6):831-837.
50. Alberio R, Croxall N, & Allegrucci C (2010) Pig epiblast stem cells depend on activin/nodal signaling for pluripotency and self-renewal. *Stem Cells Dev* 19(10):1627-1636.
51. Stern DL (2013) The genetic causes of convergent evolution. *Nature reviews* 14(11):751-764.
52. Leclère L*, et al.* (2012) Maternally localized germ plasm mRNAs and germ cell/stem cell formation in the cnidarian Clytia. *Dev. Biol.* 364(2):236-248.
53. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24(8):1586-1591.
54. Hurst LD (2002) The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* 18(9):486.
55. Felsenstein J (1985) Phylogenies and the Comparative Method. *Am. Nat.* 125(1):1-15.
56. Lanfear R, Welch JJ, & Bromham L (2010) Watching the clock: studying variation in rates of molecular evolution between species. *Trends Ecol Evol* 25(9):495-503.
57. Swanson WJ & Vacquier VD (2002) The rapid evolution of reproductive proteins. *Nat. Rev. Genet.* 3(2):137-144.
58. Singh R & Jagadeeshan S (2012) Sex and speciation: Drosophila reproductive tract proteins- twenty five years later. *Int J Evol Biol* 2012:191495.
59. Singh A & Singh BN (2014) Role of sexual selection in speciation in Drosophila. *Genetica* 142(1):23-41.
60. Edwards SV*, et al.* (2005) Speciation in birds: genes, geography, and sexual selection. *Proc Natl Acad Sci U S A* 102 Suppl 1:6550-6557.

61. Webster PJ, Suen J, & Macdonald PM (1994) *Drosophila virilis oskar* transgenes direct body patterning but not pole cell formation or maintenance of mRNA localization in *D. melanogaster*. *Development* 120(7):2027-2037.
62. Ahuja A & Extavour CG (2014) Patterns of molecular evolution of the germ line specification gene *oskar* suggest that a novel domain may contribute to functional divergence in *Drosophila*. *Dev. Genes Evol.* 222(4):65-77.
63. Ewen-Campen B, Schwager EE, & Extavour CG (2010) The molecular machinery of germ line specification. *Mol. Reprod. Dev.* 77(1):3-18.
64. Machado JP*, et al.* (2016) Bone-associated gene evolution and the origin of flight in birds. *BMC Genomics* 17:371.
65. Ducy P & Karsenty G (2000) The family of bone morphogenetic proteins. *Kidney Int.* 57(6):2207-2214.
66. Park GT & Morasso MI (2002) Bone morphogenetic protein-2 (BMP-2) transactivates Dlx3 through Smad1 and Smad4: alternative mode for Dlx3 induction in mouse keratinocytes. *Nucleic Acids Res.* 30(2):515-522.
67. Buckley D (2012) Evolution of Viviparity in Salamanders (Amphibia, Caudata). *eLS*.
68. Blackburn DG (1999) Viviparity and oviparity: Evolution and Reproductive Strategies. *Encyclopedia of Reproduction*, eds Knobil E & Neill JD (Academic Press, London), pp 994-1003.
69. Blackburn DG (2006) Squamate Reptiles as Model Organisms for the Evoluiton of Viviparity. *Herpetological Monographs* 20(1):131-146.
70. Blackburn DG (1985) Evolutionary origins of viviparity in the Reptilia. II. Serpentes, Amphisbaenia, and Ichthyosauria. *Amphibia-Reptilia* 6(3):259-291.
71. Thuesen EV (2003) Crossota millsae (Cnidaria: Trachymedusae: Rhopalonematidae), a new species of viviparous hydromedusa from the deep sea off California and Hawaii. Zootaxa. *Zootaxa* 1:1-12.
72. Mendoza G (1962) The reproductive cycles of three viviparous teleosts, Alloophorus robustus, Goodea luitpoldii and Neoophorus diazi. *Biol. Bull.* 123:351-365.
73. Hirose E (2003) Colonial allorecognition, hemolytic rejection, and viviparity in botryllid ascidians. *Zool. Sci.* 20(4):387-394.
74. Smyth JD & Wakelin D (1994) *Introduction to animal parasitology* (Cambridge University Press) 3rd Ed.
75. Lode T (2012) Oviparity or viviparity? That is the question. *Reprod Biol* 12(3):259-264.
76. Wourms JP & Lombardi J (1992) Reflections on the Evolution of Piscine Viviparity. *Integr. Comp. Biol.* 32(2):276-293.
77. Drake JW, Charlesworth B, Charlesworth D, & Crow JF (1998) Rates of spontaneous mutation. *Genetics* 148(4):1667-1686.
78. Hill KA*, et al.* (2005) Tissue-specific time courses of spontaneous mutation frequency and deviations in mutation pattern are observed in middle to late adulthood in Big Blue mice. *Environ. Mol. Mutagen.* 45(5):442-454.
79. Datta A & Jinks-Robertson S (1995) Association of increased spontaneous mutation rates with high levels of transcription in yeast. *Science* 268(5217):1616-1619.
80. Green P, Ewing B, Miller W, Thomas PJ, & Green ED (2003) Transcription-associated mutational asymmetry in mammalian evolution. *Nat. Genet.* 33(4):514-517.