Mechanisms of germ cell specification across the metazoans: epigenesis and preformation

Cassandra G. Extavour* and Michael Akam

Laboratory for Development and Evolution, University Museum of Zoology, Department of Zoology, Downing Street, Cambridge CB2 3EJ, UK

*Author for correspondence (e-mail: cgme2@hermes.cam.ac.uk)

Summary

Germ cells play a unique role in gamete production, heredity and evolution. Therefore, to understand the mechanisms that specify germ cells is a central challenge in developmental and evolutionary biology. Data from model organisms show that germ cells can be specified either by maternally inherited determinants (preformation) or by inductive signals (epigenesis). Here we review existing data on 28 metazoan phyla, which indicate that although preformation is seen in most model organisms, it is actually the less prevalent mode of germ cell specification, and that epigenetic germ cell specification may be ancestral to the Metazoa.

Supplemental data available online

Introduction

Germ cell segregation is an important problem in developmental biology, as it addresses how the fundamental distinctions between germ cells and somatic cells are initiated and maintained throughout development. The timing and mechanism of this segregation are also important for our understanding of evolution, for these influence the selective pressures that act on germ cells prior to gametogenesis, and so have important consequences for the selection of heritable variation (Extavour and García-Bellido, 2001).

Primordial germ cells of many different species share intrinsic qualities that differentiate them from somatic cells, often long before the somatic gonads are formed. However, there has been a history of disagreement as to how germ cells may be identified, and when in development the germ line is specified. In this review, we examine descriptive and experimental data on the timing and mode of origin of the germ cell lineage throughout the animal kingdom.

There are at least two distinct modes of germ line segregation in animals, both of which are well documented from experimental studies in model systems. These modes are summarised in Box 1. In some species, germ cells can easily be identified very early in embryogenesis, when their differentiation as germ cells is assured by the localisation of maternally inherited determinants before, or immediately following, fertilisation (‘preformation’). In other species, germ cells are not observed until later in development, and arise as a result of inductive signals from surrounding tissues (‘epigenesis’).

To avoid confusion, the terminology that we will use in this review for germ cells and their precursors follows the nomenclature of Nieuwkoop and Sutasurya (Nieuwkoop and Sutasurya, 1979). When germ cells become sexually differentiated and enter the first stages of gametogenesis, they are collectively termed gonia (oogonia and spermatogonia). Through the processes of oogenesis and spermatogenesis, gonia become oocytes and spermatocytes, maturing finally into ova and spermatozoa, respectively. Many organisms generate their gonia from cells capable of almost indefinite rounds of asymmetric, self-renewing mitotic divisions; these cells are called germ line stem cells. The first cells that will give rise exclusively to germ cells by clonal mitotic divisions are called primordial germ cells (PGCs). The precursors to the PGCs, which are often initially morphologically indistinguishable from the surrounding somatic cells, are called presumptive primordial germ cells (pPGCs). These divide mitotically to produce one PGC and one somatic cell.

Several aspects of germ cell morphology and function are clearly similar across many phyla of animals (Box 2). In spite of this, the mechanisms that generate germ cells appear to be highly variable, involving either prelocalised determinants or inductive processes. Previous monographs on comparative germ cell specification are now over 20 years old (Bounoure, 1939; Nieuwkoop and Sutasurya, 1979; Nieuwkoop and Sutasurya, 1981; Wolff, 1964). This review examines over 150 years worth of data on modes of germ cell specification in 28 metazoan phyla, expanding previous studies with the addition of recent molecular and experimental data. In this article we have also mapped the data onto a modern phylogeny of the Metazoa, to address the question of the ancestral mode and evolution of germ cell specification mechanisms. We conclude, in agreement with earlier surveys, that epigenesis is a more frequent mode of germ cell specification than preformation. This finding, together with data on germ cell origin in basal metazoans, suggests that epigenesis may have been the ancestral mechanism of early metazoan germ cell segregation. Our conclusion challenges a widely held view in the field of developmental biology (e.g. Wolpert, 1998) that epigenetic germ cell determination is an exception, and that most animals use localised cytoplasmic determinants to specify the germ line.

In the following sections, we review data on the earliest specification of germ cells in development, in both the
bilaterian animals (see Box 3) and their outgroups. We first consider findings in the few well-studied model organisms, and then the much wider range of studies on non-model organisms. [As we present the conclusions of the extensive studies on model organisms only briefly, we refer the reader to other reviews for further detail (Houston and King, 2000b; Matova and Cooley, 2001; Noce et al., 2001; Saffman and Lasko, 1999; Wylie, 2000)].

**Germ cell specification in model systems**

**Preformation in germ cell development**

The most comprehensive data set on the molecular mechanisms of germ cell specification is that available for *Drosophila melanogaster*. Before blastoderm formation, precocious cellularisation at the posterior pole of the embryo creates four to five pole cells (Huettner, 1923), which are the exclusive progenitors of the germ line (Fig. 1A) (Technau and Campos-Ortega, 1986; Williamson and Lehmann, 1996). The pole cells acquire PGC identity through the inheritance of specialised pole plasm, which is assembled at the posterior pole of the oocyte before fertilisation (reviewed by Mahowald, 2001). Transplantation (Illmensee and Mahowald, 1974; Illmensee and Mahowald, 1976; Illmensee et al., 1976) or forced assembly of pole plasm in ectopic sites, such as at the anterior of the oocyte (Ephrussi and Lehmann, 1992), results in PGC formation at these sites, which indicates that the pole plasm is a true germ cell determinant, and not simply a germ cell marker. Germ cell specification in *D. melanogaster* is obviously driven by preformation. In fact, it seems that all Diptera (see Box 3) localise pole plasm and form pole cells, although this preformation with respect to germ line segregation is not representative of most insects (see discussion below).

*Caenorhabditis elegans* embryos contain electron-dense granules called P granules, which are scattered evenly throughout the cytoplasm before and just after fertilisation, but which then move to the posterior of the embryo during pronuclear fusion (Hird et al., 1996). These granules are asymmetrically segregated during the unequal early cleavages so that the small P4 blastomere of the 16- to 24-cell embryo contains all of them and is the single PGC (Deppe et al., 1978; Strome and Wood, 1982). *C. elegans* thus provides a second example where germ cells are likely to be specified by preformation. In fact, it seems that all nematodes (see Box 3) localise pole plasm and form pole cells, although this preformation with respect to germ line segregation is not representative of most insects (see discussion below).

*Caenorhabditis elegans* embryos contain electron-dense granules called P granules, which are scattered evenly throughout the cytoplasm before and just after fertilisation, but which then move to the posterior of the embryo during pronuclear fusion (Hird et al., 1996). These granules are asymmetrically segregated during the unequal early cleavages so that the small P4 blastomere of the 16- to 24-cell embryo contains all of them and is the single PGC (Deppe et al., 1978; Strome and Wood, 1982). *C. elegans* thus provides a second example where germ cells are likely to be specified by preformation. In fact, it seems that all Diptera (see Box 3) localise pole plasm and form pole cells, although this preformation with respect to germ line segregation is not representative of most insects (see discussion below).

**Studies on anuran amphibian embryos** (see Box 3) have provided some of the first experimental evidence of preformation and the role of germ plasm in vertebrate germ cell specification (Bounoure, 1939). During *Xenopus laevis* oogenesis, specialised cytoplasm is synthesized and localised to the vegetal subcortex. This vegetal plasm is characterised by an accumulation of mitochondria (sometimes called the mitochondrial cloud, see Box 3) that is associated with electron-dense granules, and specific proteins and RNAs (Heasman et al., 1984; Houston and King, 2000a; Kloc et al., 2001; Kloc et al., 2002; Zhou and King, 1996). Following fertilisation, the vegetal plasm forms patchy aggregates in the vegetal hemisphere, which are segregated unequally into cleavage cells and finally accumulate specifically in a few cells that become the PGCs (Whittington and Dixon, 1975). Experiments that compromise the vegetal plasm by physical
Box 2. Germ cell identification

Germ cells can often be distinguished from somatic cells during early development using histological and molecular characteristics. Studies to define the embryonic origin of germ cells should show that putative primordial germ cells (PGCs) satisfy as many of these identification criteria as possible. In laboratory organisms, descriptive techniques can be combined with experimental methods to provide conclusive proof of PGC identity. Although experimental data are not available for most non-model organisms considered here, often a combination of histological and molecular data can indicate the site and developmental timing of PGC formation.

Histological characteristics

Until the advent of molecular techniques, most cell types were identified by their histological characteristics. Germ cells were recognised by their characteristic large round nucleus, single large nucleolus, cytoplasm relatively clear of organelles, and granular cytoplasmic material (called ‘nuage’, see below). These features are shown in the drawing of a resting and a dividing primordial germ cell from the genital ridge of the turtle *Sternotherus odoratus* (Risley, 1933). Modern studies using molecular criteria have generally confirmed PGC identifications made using older histological methods.

Electron-dense cytoplasmic bodies

Transmission electron microscopy (TEM) has revealed that electron-dense masses exist in the cytoplasm of germ cells of all phyla studied to date. These dense bodies are often called nuage or germ granules, and can be used to identify PGCs at early developmental stages. Germ cell-specific organelles (such as the mitochondrial cloud and Balbiani body) contain dense bodies.

Molecular markers

Enzyme markers, such as alkaline phosphatase, can be used to identify PGCs. However, because these markers are not always expressed by PGCs at all stages of development, they are usually only suitable for identifying germ cells at certain times of development. Modern studies often identify PGCs by identifying the products of germ cell-specific genes (see Table 1 for genes and proteins involved in germ cell specification and identity, which are useful as germ cell markers in a range of species).

Products of the vasa gene family are the most widely used molecular PGC markers. *vasa* encodes a DEAD-box RNA helicase that is usually expressed specifically in the germ line. The high conservation of motifs in these genes have made them easy to clone from many phyla. The accompanying figure shows anti-Vasa antibody staining in PGCs of the crustacean *Parhyale hawaiensis* (C.G.E., unpublished).

Transcriptional and translational regulation

When first specified, PGCs often remain transcriptionally quiescent, while the surrounding soma is usually transcriptionally active. Germ cell transcriptional repressors can be gene-specific (e.g. germ cell-less in *Drosophila*) or global (e.g. pie-1 in *C. elegans*). Translational repression in the germ line has also been documented in *Drosophila* and *C. elegans*, but it is not clear how widely the mechanisms are shared.

The origin of PGCs in the zebrafish *Danio rerio* was unclear (Lin et al., 1992; Walker and Streisinger, 1983) until the identification of a *vasa*-like gene in 1997 (Olsen et al., 1997; Yoon et al., 1997). *vasa* mRNA is synthesized during oogenesis, localises to the cleavage furrows during the first embryonic cleavages, and seems to be thereby drawn into clumps that segregate into four cells by the 32-cell stage of embryogenesis (Yoon et al., 1997). These four cells become the PGCs. Cell lineage studies and *vasa* expression patterns in other fish (Braat et al., 1999) suggest that preformation may be a common mechanism for the teleosts, but not necessarily for all fish (see Table 2, Table S2).

Chicken germ cells were thought to originate from the hypoblast (see Box 3) (Swift, 1914) until 1981, when experiments using chick-quaill chimaeras made before primitive streak formation showed that they were of epiblastic origin (Eyal-Giladi et al., 1981). PGCs were then thought to arise through a gradual epigenetic process beginning at around stage X (an intrauterine early blastoderm stage) (Karagenc et al., 1996; Naito et al., 2001). However, the recent isolation of a chicken *vasa* homologue has made it possible to trace pPGCs as far back as cleavage stage embryos (Tsunekawa et al., 2000). Chicken vasa protein forms part of the mitochondrial cloud in chick oocytes, and localises to cleavage furrows until stage IV, when six to eight cells of the ~300-cell embryo contain vasa and are good candidates for the PGCs (Tsunekawa et al., 2000). These data suggest that preformation may be the mechanism for germ cell specification in chickens, although functional studies have yet to be carried out.
Box 3. Glossary of terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>4d CELL</td>
<td>The mesentoblast cell of spirally cleaving animals; gives rise to both mesoderm and endoderm.</td>
</tr>
<tr>
<td>ALLANTOIS</td>
<td>A mesoderm-derived structure that emerges from the posterior end of the embryo and attaches to the placenta. It gives rise to the placental blood vessels and the umbilical cord.</td>
</tr>
<tr>
<td>AGNATHA</td>
<td>A grade of chordate, including hagfish and lampreys, characterized by the absence of jaws.</td>
</tr>
<tr>
<td>ANURA</td>
<td>Amphibian order that includes those without a tail, such as frogs and toads.</td>
</tr>
<tr>
<td>BALBIANI BODY</td>
<td>Found in oocytes of some species, this organelle contains mitochondria, Golgi vesicles, centrosomes and endoplasmic reticulum; also called the yolk nucleus or vitelline body; probably a condensed form of the mitochondrial cloud.</td>
</tr>
<tr>
<td>BASAL</td>
<td>An evolutionary lineage, or animal within a lineage, that arises close to the root or base within a phylogeny.</td>
</tr>
<tr>
<td>BILATERIA</td>
<td>Animals that show bilateral symmetry across a body axis.</td>
</tr>
<tr>
<td>CHAETOGNATHA</td>
<td>The phylum of arrow worms, small transparent marine worms found both in the plankton and in the benthos.</td>
</tr>
<tr>
<td>CLADE</td>
<td>A lineage of organisms that comprises an ancestor and all its descendants.</td>
</tr>
<tr>
<td>COLLEMBOLA</td>
<td>The arthropod order of direct-developing, wingless hexapods, also known as springtails.</td>
</tr>
<tr>
<td>DERIVED</td>
<td>Evolved to a state that is not like the primitive condition.</td>
</tr>
<tr>
<td>DEUTEROSTOME</td>
<td>A bilaterian animal whose mouth forms as a secondary opening, separate from the blastopore.</td>
</tr>
<tr>
<td>DIPLOBLAST</td>
<td>Animals with only two germ layers (ectoderm and endoderm), including the Cnidaria and Ctenophora.</td>
</tr>
<tr>
<td>DIPNOI</td>
<td>The subclass of sarcopterygian fishes known as lungfishes, which breathe by a modified air bladder, as well as gills.</td>
</tr>
<tr>
<td>DIPTERA</td>
<td>The insect order of true flies that bear only one pair of functional wings, such as Drosophila melanogaster, mosquitoes, gnats and midges.</td>
</tr>
<tr>
<td>ECDYSOZOA</td>
<td>A protostome clade of moulting animals that includes both C. elegans and D. melanogaster, but not annelids.</td>
</tr>
<tr>
<td>ENTEROPNEUSTA</td>
<td>The subphylum of hemichordates known as the acorn worms.</td>
</tr>
<tr>
<td>EPIBLAST</td>
<td>The embryonic layer of vertebrate embryos from which the embryo proper arises during gastrulation; gives rise to all three germ layers of the embryo.</td>
</tr>
<tr>
<td>HOMOLOGOUS</td>
<td>A character in two or more taxa with a unique origin in the common evolutionary ancestor of those taxa. A statement of homology is an evolutionary hypothesis, and relates to a particular attribute of a structure or process. For further discussion, see Bolker and Raff (Bolker and Raff, 1996).</td>
</tr>
<tr>
<td>HOMOLOGOUS BODY</td>
<td>Found in oocytes of some species, this organelle contains mitochondria, Golgi vesicles, centrosomes and endoplasmic reticulum; also called the yolk nucleus or vitelline body; probably a condensed form of the mitochondrial cloud.</td>
</tr>
<tr>
<td>HYPOBLAST</td>
<td>Older term for the inner germ layer in bird and reptile embryos; the origin of the endoderm.</td>
</tr>
<tr>
<td>LOPHOTROCHOZOA</td>
<td>A clade of protostomes supported by most molecular phylogenies, including spirally cleaving animals such as molluscs and annelids, as well as lophophorates such as brachiopods and phoronids.</td>
</tr>
<tr>
<td>METATHERIA</td>
<td>Marsupials: mammals that give birth to live offspring and suckle young in maternal pouches.</td>
</tr>
<tr>
<td>MITOCHONDRIAL CLOUD</td>
<td>An organelle composed of a high concentration of mitochondria, containing electron dense cytoplasm similar to germ plasm. Probably a diffuse form of the Balbiani body.</td>
</tr>
<tr>
<td>MONOTREMATA</td>
<td>The egg-laying mammals (platypuses and echidnas).</td>
</tr>
<tr>
<td>OOPLASM</td>
<td>The cytoplasm of the oocyte or unfertilised egg.</td>
</tr>
<tr>
<td>PROTOSTOME</td>
<td>A bilaterian animal whose mouth and anus develop from the same invagination (the blastopore) during embryogenesis.</td>
</tr>
<tr>
<td>PHYLUM</td>
<td>The highest taxonomic category used to subdivide the animals or species of any other taxonomic kingdom.</td>
</tr>
<tr>
<td>SARCOPTERYGII</td>
<td>The vertebrate group that includes lobe-finned fish and tetrapods, including lungfishes and coelacanths.</td>
</tr>
<tr>
<td>SAUROPSIDA</td>
<td>A group of vertebrates including birds, dinosaurs and reptiles other than turtles.</td>
</tr>
<tr>
<td>TRIPLOBLAST</td>
<td>An animal with three germ layers (ectoderm, mesoderm and endoderm).</td>
</tr>
<tr>
<td>URODELA</td>
<td>An order of amphibians including axolotls, salamanders and newts.</td>
</tr>
</tbody>
</table>

Epigenesis in germ cell development

The time and site of origin of mammalian germ cells was a controversial issue for several decades (see Everett, 1945; Heys, 1931) (see also references in Table S2 online). This controversy continued until alkaline phosphatase activity was first used in 1954 as a marker for mouse germ cells (Chiquoine, 1954). This technique was later used to identify these cells in mouse embryos between 7 and 7.5 days post coitum (dpc) (Fig. 1B) (Ginsburg et al., 1990; Ozdzenski, 1967). In 1994, lineage tracing studies moved the time of origin of these cells to an even earlier stage of development, 6.5 dpc (Lawson and Hage, 1994). At this stage of development, these cells are found posterior to the primitive streak in the extraembryonic mesoderm, at the base of the allantois (see Box 3). They are incorporated into the hindgut epithelium, move into the dorsal mesentery, and from there, they colonise the genital ridges on the dorsal body wall, forming the gonad primordia (Chiquoine, 1954; Ginsburg et al., 1990; Gomperts et al., 1994).

In contrast to the studies in chick and zebrafish, the isolation of a mouse vasa homologue has not resulted in the identification of pPGCs at even earlier stages of development (Fujiwara et al., 1994; Noce et al., 2001; Toyooka et al., 2000). Although mouse vasa homologue protein is expressed in oocytes, it is not localised to a specific subcellular region, and no germ plasm is formed (Toyooka et al., 2000). Instead, a true epigenetic mechanism for germ line specification has been demonstrated by both descriptive and experimental evidence (Tsang et al., 2001). Cells of the distal epiblast (see Box 3), which normally differentiate into ectodermal derivatives, can differentiate as PGCs when transplanted into the proximal epiblast, the region from which the PGCs normally derive. Conversely, proximal epiblast cells will not differentiate as PGCs when transplanted to distal sites (Tam and Zhou, 1996). These experiments suggested that inductive signals might be required for germ cell specification in the mouse. At least some of these inductive signals have been identified as members of the bone morphogenetic protein (BMP) class of TGFβ superfamily intercellular signaling proteins (Hogan, 1996).

The expression of Bmp4 (Lawson et al., 1999) and Bmp8b (Ying et al., 2000) in the extraembryonic ectoderm, and Bmp2 in the endoderm (Ying and Zhao, 2001), is required for the induction of germ cell fate among proximal epiblast cells. A study of gene expression at the single cell level has indicated that the genes fragilis and stella are upregulated in a subset of the proximal epiblast cells. The expression of these two genes appears to make the cells competent to respond to BMP
signals, which direct them to differentiate into PGCs (Saitou et al., 2002). However, even cells of the distal epiblast, which do not normally express fragilis or stella, can be induced to differentiate into PGCs if placed next to the source of the BMP signals (Tam and Zhou, 1996). These results tell us that germ cell specification in mice is clearly epigenetic and does not depend on maternally localised determinants.

The only other unequivocal evidence for inductive germ cell specification has arisen from studies on urodele amphibians (see Box 3). Germ cells were first identified in the lateral plate mesoderm (LPM) of many urodele species (Humphrey, 1925; Humphrey, 1929; Ikenishi and Nieuwkoop, 1978). Careful explant and grafting experiments have shown that the LPM is not merely the place in which germ cells could first be unambiguously identified, but also that these cells actually arose there as a result of inductive signals from the ventral endoderm (Boterenbrood and Nieuwkoop, 1973; Nieuwkoop, 1947). These signals induce both PGCs and other somatic cell types. Kotani showed that presumptive epidermal cells placed at the site of the LPM can give rise to PGCs (Kotani, 1957), and later studies demonstrated that any part of the animal half of the blastula can give rise to PGCs under the inductive influence of the ventral endoderm (Sutasurya and Nieuwkoop, 1974). Recent studies in the axolotl Ambystoma mexicanum have confirmed that both a mitochondrial cloud and localised molecular determinants are absent in oocytes of this organism (Johnson et al., 2001) (A. D. Johnson, M. Drum and R. Bachvarova, unpublished). The products of germ cell-specific genes, such as Dazl and vasa, are not localised in the oocytes or early embryos of this axolotl, and are not zygotically transcribed in PGCs until they approach the gonadal ridges (Johnson et al., 2001; Johnson et al., 2003). Although no data are available yet on the molecular nature of the endodermal signal that induces PGC and LPM differentiation in urodeles, BMP4 is known to induce ventral mesoderm in X. laevis (Dale et al., 1992; Jones et al., 1992), and it is therefore possible that this signal plays a role in axolotl PGC specification.

Germ cell specification in non-model systems

The laboratory models we have considered thus far are members of only three bilaterian phyla (Arthropoda, Nematoda and Chordata), and cannot be considered to represent the diversity of the Metazoa. To evaluate the distribution of preformation and epigenesis as modes of germ cell specification, we now summarise what is known about the mode of germ cell formation for each metazoan phylum. Most of these data do not provide conclusive evidence, but the fulfillment of multiple criteria for the identification of PGCs, together with experimental evidence, strongly indicate the mode of PGC determination in many such phyla.

In Table 2, we present recent molecular data, and older descriptive and experimental literature, on comparative germ cell specification (for a fully referenced version of this table, see Table S2 online at http://dev.biologists.org/supplemental/). This table lists the phyla that we have reviewed; the observed location, developmental timing and presumed mode of germ cell specification; and whether functional experiments have been carried out to distinguish between epigenesis and preformation. The criteria used to identify pPGCs and PGCs are also indicated. In the following section, specific references are given only for a few examples of each major clade (see Box 3); references for all other statements can be found in Table S2 at http://dev.biologists.org/supplemental/.

Origin of germ cells in basal animal lineages

Porifera (sponges) and Cnidaria (corals, jellyfish, hydra) are the most basal (see Box 3) branches of the Metazoa. In these phyla, germ cells arise from a stem cell population that also generates other cell types. Thus, the boundary between germ line and soma is a fluid one. For this reason, these basal groups are sometimes omitted from comparative discussions of germ cell origin (e.g. Dixon, 1994; Ransick et al., 1996). However, these organisms can produce haploid gametes and reproduce sexually, and in that sense their germ line serves the same function as it does in bilaterian animals. In hydrozoan cnidarians, pluripotent cells called interstitial cells (I cells) contain electron-dense cytoplasmic bodies similar to those associated with germ cells in all phyla (Eddy, 1975). These bodies become more numerous in I cells that develop into germ cells, and decrease in number in I cells that differentiate into nematocytes. In Porifera, archaeocytes are pluripotent cells that are capable of both germ line and somatic stem cell divisions.

Ctenophores (comb jellies) also probably diverged from other Metazoa before the origin of the Bilateria (Fig. 2A). Ctenophore germ cells have been described as arising epigenetically, from the meridional canal endoderm (Fig. 2B), but their having an extragonadal origin, followed by their migration to the meridional canal primordium, cannot be ruled out.

Fig. 1. Germ cell specification in model systems. (A) A cellular blastoderm stage D. melanogaster embryo stained with anti-Vasa antibody. The pole cells (arrowhead), located at the posterior pole of the embryo, are the primordial germ cells (PGCs), and express vasa protein. (B) Mouse embryo at 7 dpc stained with alkaline phosphatase. Enzymatic activity is high in the PGCs (arrowhead), which are located in the proximal epiblast at the base of the allantois [Reproduced with permission from McLaren (McLaren, 2003)]. Anterior is to the left in both panels.
<table>
<thead>
<tr>
<th>Gene (common name)</th>
<th>Fly (D)</th>
<th>Worm (C)</th>
<th>Frog (X)</th>
<th>Fish (Dr)</th>
<th>Mouse (M)</th>
<th>Other‡</th>
<th>Gene product</th>
<th>Germ cell function§</th>
</tr>
</thead>
<tbody>
<tr>
<td>boule</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td>RNP-type RNA binding protein with DAZ repeats</td>
<td>Meiosis; PGC differentiation (Hs, M, X)</td>
</tr>
<tr>
<td>aubergine</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Similar to eIF2C (translation initiation factor)</td>
<td>Pole cell formation; translational regulation of osk</td>
</tr>
<tr>
<td>bruno</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td>Hs</td>
<td>RNP-type binding domains</td>
<td>Translational regulation of osk and grk (D)</td>
</tr>
<tr>
<td>capuccino</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Actin binding protein</td>
<td>osk &amp; stau localisation in oocyte (D)</td>
</tr>
<tr>
<td>DEADSouth</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>eIF4A-like helicase</td>
<td>Localised to germ granules (X)</td>
</tr>
<tr>
<td>fragilis</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IFN inducible TM family member</td>
<td>Confers PGC competence (M)</td>
</tr>
<tr>
<td>germ-cell-less</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td>Nuclear pore associated protein</td>
<td>Transcriptional repression (D)</td>
</tr>
<tr>
<td>gld-1</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KH motif RNA binding protein</td>
<td>Mutant has fewer PGCs (M)</td>
</tr>
<tr>
<td>gp130</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cytokine receptor</td>
<td>EGFR ligand</td>
</tr>
<tr>
<td>gurken</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>Oocyte patterning and germ plasm assembly (D)</td>
</tr>
<tr>
<td>gustavus</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RNA-dependent ATPase</td>
<td>G plasm component localisation (D)</td>
</tr>
<tr>
<td>homeless</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>Germ plasm assembly (C, D)</td>
</tr>
<tr>
<td>mago nashi</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>Hs</td>
<td>Similar to E(z) (D polycomb gene)</td>
<td>Transcriptional repression (C)</td>
</tr>
<tr>
<td>mes-2</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>MES-2 and MES-6 localisation (C)</td>
</tr>
<tr>
<td>mes-3</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>GC survival (C)</td>
</tr>
<tr>
<td>mes-4</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>Transcriptional repression, MES-2 localisation (C)</td>
</tr>
<tr>
<td>mes-6</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zinc finger protein</td>
<td>PIE-1 and P granule segregation (C)</td>
</tr>
<tr>
<td>mex-1</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KN domain RNA binding protein</td>
<td>Blastomere identity; mutation leads to ectopic GCs (C)</td>
</tr>
<tr>
<td>mex-3</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mitochondrial ribosomal RNA</td>
<td>Localisation of mitochondrial ribosomes on P granules (D)</td>
</tr>
<tr>
<td>mtlrRNA</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>Germ plasm assembly (C, D)</td>
</tr>
<tr>
<td>nanos</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>Ch, Dv, Gd, H (Comiss, Comox2), Hr (Hrnos), S, Md</td>
<td>CCHC Zn-finger protein</td>
<td>Translational and transcriptional repression (C, Ch, D, Dv, Md)</td>
</tr>
<tr>
<td>orb</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RNA binding protein</td>
<td>osk localisation (D)</td>
</tr>
<tr>
<td>oskar</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td>Dv</td>
<td>Novel protein</td>
<td>Germ plasm assembly (D)</td>
</tr>
<tr>
<td>par-1</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td>Hs, R</td>
<td>Novel protein</td>
<td>OSK phosphorylation, germ plasm assembly (C, D)</td>
</tr>
<tr>
<td>pge-1</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Non-coding RNA</td>
<td>PC migration (D)</td>
</tr>
<tr>
<td>pie-1</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zinc finger protein</td>
<td>Transcriptional repression (C)</td>
</tr>
<tr>
<td>pOG</td>
<td>yes</td>
<td></td>
<td>yes</td>
<td></td>
<td></td>
<td>Hs (CUG-BP) S</td>
<td>Plant homeodomain motifs</td>
<td>PGC proliferation (M)</td>
</tr>
<tr>
<td>pumilio</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td>Hs</td>
<td>Novel RNA binding domains</td>
<td>Transcriptional repression (D, C)</td>
</tr>
<tr>
<td>spire</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>osk and stau localisation in oocyte (D)</td>
</tr>
<tr>
<td>staufen</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>dsRNA binding protein</td>
<td>Germ plasm assembly (D)</td>
</tr>
<tr>
<td>stella</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>Confers PGC competence (M)</td>
</tr>
<tr>
<td>tropomyosin II</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Actin binding protein</td>
<td>osk and stau localisation in oocyte (D)</td>
</tr>
</tbody>
</table>

*Table 1. Genes required by germ cells for development*
Only one other group of animals is now thought to have diverged from the bilaterian stem before the split between protostomes and deuterostomes. These are the accoelomorph flatworms (accoels and nemertodermatids) (Ruiz-Trillo et al., 2002; Telford et al., 2003). Several molecular datasets suggest that they are basal to the Bilateria, and not closely related to the other flatworms in the phylum Platyhelminthes. Germ cells in accoels are derived from a population of pluripotent cells and are the cells that make regeneration possible in these animals. There is no evidence for germ line determination by preformation in any of these basal animal lineages.

**Germ cell specification in bilaterian animals**

Recent metazoan phylogenies based on molecular characters suggest that, with the exception of the basal animal groups mentioned above, all animals fall within one of three great lineages, each of which includes both simple and complex animals (Adoutte et al., 2000; Peterson and Eernisse, 2001). These three clades are the deuterostomes (which include the chordates), and two clades of protostomes, the ecdysozoans (which include C. elegans and Drosophila) and the lophotrochozoans (for which there are no well-studied laboratory models) (see Box 3). In Table 2, the phyla are organised into these groupings, although in the text we consider the protostomes as a whole (see supplemental Data 1 at http://dev.biologists.org/supplemental/ for a guide to the taxonomic groupings used in Table 2). Within each of these clades, the relationships between phyla are poorly resolved, so at present it is not easy to predict which phyla are most likely to retain ancestral characteristics.

**PGCs in protostomes**

Drosophila and C. elegans developmental studies have provided us with so much molecular genetic information on germ cell specification that it is easy to forget how little is known about the other protostomes, which include at least 20 phyla and make up the vast majority of animal species (Brusca and Brusca, 2003). A few remarkable cases of germ plasm segregation have indeed been observed outside of fruit flies and nematode worms. For example, fertilized eggs of the bivalve mollusc Sphaerium striatum contain an asymmetrically localised dense matter, which is segregated during unequal cleavages to the 4d cell (Woods, 1931; Woods, 1932). The 4d localised dense matter, which is segregated during unequal cleavages to the 4d cell (Woods, 1931; Woods, 1932). The 4d localised dense matter, which is segregated during unequal cleavages to the 4d cell (Woods, 1931; Woods, 1932).

---

### Table 1. Continued

<table>
<thead>
<tr>
<th>Gene (common name)</th>
<th>Fly (D)</th>
<th>Worm (C)</th>
<th>Frog (X)</th>
<th>Fish (Dr)</th>
<th>Mouse (M)</th>
<th>Other</th>
<th>Gene product</th>
<th>Germ cell function</th>
</tr>
</thead>
<tbody>
<tr>
<td>tudor</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel ‘tudor domain’ repeats</td>
<td>Germ plasm assembly; nos localisation (D)</td>
</tr>
<tr>
<td>valois</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>Germ plasm assembly (D)</td>
</tr>
<tr>
<td>vasa</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>**</td>
<td>DEAD-box RNA helicase; eIF4A (translation initiation factor) homology</td>
<td>Germ plasm assembly; translational regulation (D)</td>
</tr>
<tr>
<td>Xlsirts</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Non-coding RNA</td>
<td>mRNA localisation to vegetal cortex (X)</td>
</tr>
<tr>
<td>Xpat</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Novel protein</td>
<td>Localised to germ plasm (X)</td>
</tr>
</tbody>
</table>

*Data compiled from 143 references, which are available in the online version of this table (see Table S1 at http://dev.biologists.org/supplemental/).

1 Usualy the name of the first gene in the family to be identifed.

2Abbreviations for species names are as follows: A, Anmbystoma mexicanum (axolotl); Ac, Aurelia aurita (moon jellyfish); Ad, Acropora digitifera (staghorn coral); B, Bombyx mori (silkworm); C, Caenorhabditis elegans (nematode); Ca, Carassius auratus (goldfish); Cb, Cebus sp. (capuchin monkey); Cc, Cyprinus carpio (carp); Ci, Chironomus samoensis (midge); Cj, Ciona intestinalis (ascidian); Cp, CnVAS1, Kunvas1, Plvas1, Vlas1, (common name) † (D) (C) (X) (Dr) (M) Other ‡ Gene product Germ cell function ¶ |

3Note that many homologues are not given new names, but may be called ‘x-like gene’, where ‘x’ is the name of the first gene in the family to be identified.

4Species for which functional information is available are in parentheses.

**AAa, Ad, B, Ca, Cc, Cc (CDEAD1), Cp, Cc, Cs (CDEAD1a, CDEAD1b), Dd (Plv31), Dj (Djvlga, Djvlgb), Dv, E (PoVAS1), Ec, G (Cv), H (CnVAS1, CnVAS2), He, Hs, Hy, L, Mf, O (olvas), On, P, R (RVLG), Sa, Sm, Sp, Sq, T, Tt.**
<table>
<thead>
<tr>
<th>Basal Lineages</th>
<th>PGC origin&lt;sup&gt;†&lt;/sup&gt;</th>
<th>Stage</th>
<th>Location/derivation</th>
<th>Mode of PGC specification&lt;sup&gt;‡&lt;/sup&gt;</th>
<th>Experimental evidence&lt;sup&gt;§&lt;/sup&gt;</th>
<th>PGC identification criteria&lt;sup&gt;‡&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porifera</td>
<td>Gastrulation</td>
<td>Mesenchymal cells</td>
<td>E</td>
<td>–</td>
<td>LM, TEM, MM</td>
<td></td>
</tr>
<tr>
<td>Cnidaria</td>
<td>Post-embryonic</td>
<td>In coelom from gastradmal cells of mesentery or endodermic epithelial cells</td>
<td>E</td>
<td>–</td>
<td>TEM, LM</td>
<td></td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Post-embryonic</td>
<td>Within ovaries from endodermally derived gastrodermis</td>
<td>E</td>
<td>–</td>
<td>TEM</td>
<td></td>
</tr>
<tr>
<td>Hydrozoa</td>
<td>Gastrulation</td>
<td>Endodermal core</td>
<td>E</td>
<td>+</td>
<td>LM, TEM, MM</td>
<td></td>
</tr>
<tr>
<td>Ceratomorpha</td>
<td>Early larval stage</td>
<td>Endoderm</td>
<td>E</td>
<td>–</td>
<td>LM</td>
<td></td>
</tr>
<tr>
<td>Bilateria</td>
<td>Lophotrochozoa (Protostomes)</td>
<td>Mesenchymal</td>
<td>E</td>
<td>–</td>
<td>LM, TEM</td>
<td></td>
</tr>
<tr>
<td>Platyhelminthes</td>
<td>Turbellaria</td>
<td>Late embryogenesis</td>
<td>Mesenchymal</td>
<td>E</td>
<td>+</td>
<td>LM, TEM, MM, MM</td>
</tr>
<tr>
<td></td>
<td>Trematoda</td>
<td>First cleavage</td>
<td>First cleavage</td>
<td>P</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Cestoda</td>
<td>Late embryogenesis</td>
<td>Mesenchymal</td>
<td>E</td>
<td>–</td>
<td>LM, TEM</td>
</tr>
<tr>
<td></td>
<td>Rotifera</td>
<td>Before gastrulation</td>
<td>4d cell</td>
<td>P</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Entoprocta</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td></td>
<td>Ectoprocta</td>
<td>Post-embryonic</td>
<td>Mesenchyme: gonadal epithelium</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Nemertea</td>
<td>Late embryogenesis</td>
<td>Mesodermally derived cells of parenchyma or gonadal epithelium</td>
<td>E</td>
<td>–</td>
<td>LM, TEM</td>
</tr>
<tr>
<td></td>
<td>Phoronida</td>
<td>Late embryogenesis</td>
<td>Peritoneal epithelium</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Brachiopoda</td>
<td>Late embryogenesis</td>
<td>Ileo-parietal epithelium</td>
<td>E</td>
<td>–</td>
<td>LM, TEM</td>
</tr>
<tr>
<td></td>
<td>Gnathostomulida</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>–</td>
<td>nd</td>
</tr>
<tr>
<td></td>
<td>Pogonophora</td>
<td>Post-embryonic</td>
<td>Gonadal epithelium</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Echiura</td>
<td>Larval stage</td>
<td>Mesoderm</td>
<td>E</td>
<td>–</td>
<td>LM, TEM</td>
</tr>
<tr>
<td></td>
<td>Sigunulida</td>
<td>Larval stage</td>
<td>Gonadal epithelium</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td>Mollusca</td>
<td>Aplacophora</td>
<td>Post-larval</td>
<td>Mesoderm?</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Polyplacophora</td>
<td>Post-embryonic</td>
<td>Gonadal epithelium</td>
<td>E</td>
<td>–</td>
<td>TEM</td>
</tr>
<tr>
<td></td>
<td>Cephalopoda</td>
<td>Blastoderm stage</td>
<td>Blastoderm superficial layer</td>
<td>P</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Gastropoda</td>
<td>Late embryogenesis/early cleavage</td>
<td>Mesoderm/early cleavage blastomere?</td>
<td>E/P</td>
<td>–</td>
<td>LM, TEM</td>
</tr>
<tr>
<td></td>
<td>Bivalvia</td>
<td>Early cleavage</td>
<td>4d cell</td>
<td>P</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td>Annelida</td>
<td>Polychaeta</td>
<td>Early cleavage/post-larval</td>
<td>E/P</td>
<td>+</td>
<td>LM, TEM</td>
</tr>
<tr>
<td></td>
<td>Oligochaeta</td>
<td>Early cleavage/late embryogenesis</td>
<td>4d cell/unknown source before mesoderm formation/unknown source late in development</td>
<td>E/P</td>
<td>+</td>
<td>LM, TEM</td>
</tr>
<tr>
<td>Hirudinea</td>
<td>Early cleavage</td>
<td>D blastomere</td>
<td>P</td>
<td>–</td>
<td>LM, MM</td>
<td></td>
</tr>
<tr>
<td>Ecdysozoa</td>
<td><em>Arthropoda</em></td>
<td>Collembola</td>
<td>Early cleavage blastomes</td>
<td>P</td>
<td>–</td>
<td>LM, TEM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insecta</td>
<td>Early cleavage</td>
<td>E/P</td>
<td>+</td>
<td>LM, TEM, SEM, EM, MM, LI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crustacea</td>
<td>Early cleavage</td>
<td>E/P</td>
<td>–</td>
<td>LM, TEM, MM, LI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chelicerata</td>
<td>Early cleavage</td>
<td>E/P</td>
<td>–</td>
<td>LM, TEM, SEM, MM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myriapoda</td>
<td>Late embryogenesis/mesoderm</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tardigrada</td>
<td>Late embryogenesis/mesoderm</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Onychophora</td>
<td>Gastrulation/late embryogenesis</td>
<td>E/P</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nematoda</td>
<td>First cleavage</td>
<td>P</td>
<td>+</td>
<td>LM, TEM, SEM, MM, LI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Priapula</td>
<td>nd</td>
<td>nd</td>
<td>–</td>
<td>nd</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gastrotricha</td>
<td>Late embryogenesis/mesoderm</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kinorhyncha</td>
<td>nd</td>
<td>Apical cells of gonad</td>
<td>E</td>
<td>–</td>
</tr>
<tr>
<td>Deuterostomes</td>
<td>Chaetognatha</td>
<td>First cleavage</td>
<td>First cleavage blastomere</td>
<td>P</td>
<td>+</td>
<td>LM, TEM, MM, LI</td>
</tr>
<tr>
<td></td>
<td>Hemichordata</td>
<td>Late embryogenesis</td>
<td>Ectoderm/mesoderm</td>
<td>E</td>
<td>–</td>
<td>LM</td>
</tr>
</tbody>
</table>
Among the hexapods, most of the basal insect orders for which data are available do not appear to have early segregated germ cells (Fig. 2C,D) (e.g. Heymons, 1891). The collemboles (see Box 3) are an exception, showing segregation of electron dense granules to PGCs in early embryonic cleavages (Klag, 1982; Klag and Swiatek, 1999; Tamarelle, 1979), but these animals may not be closely related to other hexapods (Nardi et al., 2003). Clear examples of preformation are generally found in some, but not all, species of higher insect orders such as Diptera (flies) (e.g. Lassman, 1936), Lepidoptera (moths and butterflies) (e.g. Berg and Gassner, 1978) and Hymenoptera (ants, bees and wasps) (e.g. Gatenby, 1917). The PGCs of most crustaceans appear to form late in development from the mesodermal cells of the coelomic cavities, although early segregation has been observed in some copepods (Fig. 2E,F) (Amma, 1911) and cladocerans (Kühn, 1913). Various authors have claimed that in some members of the chelicerate order Arachnida, the PGCs are segregated early in embryogenesis, forming a clump of cells between the yolk and the embryonic primordium (e.g. Juberthie, 1964). However, most studies of both chelicerate and myriapod embryogenesis show no evidence for early segregated cytoplasmic determinants, and instead report a late mesodermal origin of PGCs (e.g. Heymons, 1901; Kautzsch, 1910). In summary, it is not at all clear what the ancestral mechanism of arthropod germ line specification might have been, but epigenesis appears to be more frequent than preformation.

Nematodes are the only protostome phylum in which all members that have been studied exhibit preformation in PGC development; all other cases of preformation in the protostomes have been observed in only a few derived species within phyla for which epigenesis is prevalent and likely ancestral. For example, most species of the Platyhelminthes derive their germ cells from neoblasts, a pluripotent cell type that gives rise to different types of somatic cells, as well as to germ cells (Gustafsson, 1976; Ladurner et al., 2000). However, the trematode flatworms are a derived group within the Platyhelminthes that segregate their germ cells by preformation at the beginning of embryogenesis (Bednarz, 1973).

Most other protostomes develop their germ cells from a subpopulation of mesodermal cells at an advanced stage of embryogenesis during the differentiation of specialised mesodermal cell types. Among lophotrochozoan protostomes with canonical spiral cleavage (such as some molluscs and some annelids), this mesodermal subpopulation is derived from one of the products of the division of the 4d mesendoblast cell. With only two documented exceptions (among the molluscs) (Dohmen and Lok, 1975; Dohmen and Verdronk, 1974; Verdronk, 1973), no putative cytoplasmic determinants have been observed in precursors of this cell, hence there is currently no evidence for preformation in most annelids and molluscs. The germ line in other groups (such as nemerteans, brachiopods and some arthropods) develops during larval stages, or continuously throughout adult development, from the mesodernally derived cells of the gonadal epithelium.

The phylogenetic position of chaetognaths (see Box 3) has been contested for many decades. Because recent studies have questioned their traditional classification as deuterostomes...
Fig. 2. Examples of germ cell identification in non-model systems.
(A) An adult Mnemiopsis leidyi (a comb jelly of the phylum Ctenophora) in which germ cells seem to arise by epigenesis. These hermaphrodites have eight rows of gonads, each gonad containing an ovary and a testis behind each of the eight rows of comb plates (also called ctene rows) (asterisks). (B) Close up of area boxed in A. Germ cells are first identified in ctenophores after the larvae hatch, next to the meridional canals that give rise to the ctene rows. Multiple ovaries (black arrowheads) and testes (white arrowheads) develop on either side of the canals. In this panel, eggs (asterisks) are being extruded through the gonducts. (C) Juvenile Blatta germanica cockroach (phylum Arthropoda). Germ cells in these insects do not appear to be determined by preformation. (D) The embryonic rudiment of B. germanica forms on the surface of the yolk (yellow). (D, part i) Germ cells (gz) are first identified at the posterior of the germ band, after formation of the mesoderm (ms). (ii) As development proceeds, germ cells continue to arise from the mesoderm of the coelomic sacs (c), which are being formed in each segment in an anteroposterior progression. (iii) The number of germ cells increases, and they populate the coelomic sacs of the segments from which the gonad will form. c, coelomic sac; ek, ectoderm; gz, germ or reproductive cells (genitalzellen); ms, mesoderm; st, stomodaeum (Heymons, 1891). (E) A copepod of the genus Cyclops (phylum Arthropoda). All copepods that have been studied segregate germ cells by preformation. (F) Embryonic cleavages of Cyclops fuscus are holoblastic and equal. (i) In the first cleavage, dense granular material associates with only one of the blastomeres (orange). (ii) The resulting two-cell stage has the granular material in only one of the blastomeres (orange). (iii) The granular material continues to be asymmetrically segregated to a single blastomere (orange) in subsequent cleavages. (iv) At the time of gastrulation, the cell containing the granular cytoplasm has divided to give rise to two cells that are located at the tip of the invaginating archenteron, which are the PGCs (Ug: Urgeschl echtszellen) (Amma, 1911). (G) Late stage embryo of the turtle Trachemys scripta (phylum Chordata), stained with Alcian Blue for cartilage and Alizarin Red for bone. Reptiles seem to segregate germ cells epigenetically. (H, part i) Section of an embryo of the turtle Sternotherus odoratus at the three somite stage. Germ cells are first identified at this stage of development, in two zones (Z) lateral to the neural groove (NG). (ii) Close up of area boxed in i. PGCs in the germ cell zone (Z) are distinguishable from somatic cells of the ventral ectoderm (VE) as large cells with round nuclei and granular cytoplasm. AC, amnion and chorion; DE, definitive endoderm; N, notochord; NG, neural groove; VE, vitelline endoderm; Z, germ cell zone. Reproduced with permission from Risley (Risley, 1933). Scale bars: 3 mm in A; 250 μm in B; 150 μm in C; 50 μm in E; 5 mm in G.
(Matus et al., 2002; Telford and Holland, 1993), we consider
them here along with the other protostomes. Elpatievsky
was the first to recognise that, in chaetognaths, a specific
cytoplasmic structure was assembled after fertilisation and
asymmetrically segregated into a single cell at the 32-cell
stage. He traced the fate of this cell and found that its four
descendants were the PGCs of the juvenile gonad (Elpatievsky,
1909). Blastomere ablation and cytoplasmic disruption
experiments (Ghirardelli, 1954; Ghirardelli, 1955), combined
with recent data showing that this cytoplasmic structure
contains a Vasa-like protein, support the idea that it may be not
only a marker, but also a determinant of germ cells (Carré et
al., 2002). Given these findings, there is little doubt that the
mode of chaetognath PGC specification is preformation.

Protostomes are a hugely diverse group, with few shared
embryological characteristics (Nielsen, 2001). However, the
present survey suggests that most protostomes use epigenesis
to specify germ cells. Preformation appears in few groups, but
was unlikely to have been used to specify the germ cells of
the last common ancestor of either ecdysozoans or
lophotrochozoans. We therefore suggest that the ancestral
protostomian mechanism for germ line specification was
epigenetic, and that germ plasm specification by preformation
evolved as a derived character several times in diverse groups.

PGCs in non-chordate deuterostomes

The deuterostomes include three major phyla, the
echinoderms, the hemichordates and the chordates. In the non-
chordate deuterostome phyla, modes of germ cell specification
are hard to classify. The only studies available on
hemichordates are early histological analyses of enteroneust
(see Box 3) development, and opinion was divided among
those researchers as to whether the PGCs were of mesodermal
or ectodermal origin (Bateson, 1885; Morgan, 1894; Spengel,
1893). There is no suggestion that PGCs are specified early in
this group.

Echinoderm gonia are presumed to originate epigenetically
from the gonadal epithelium in juveniles and throughout adult
life. Regeneration of PGCs, presumably from mesenchymal
cells, has been observed even in fragments of animals without
gonads. The small micromeres of the 16-cell echinoid embryo
seem to share some mitotic characteristics with dipteran pole
cells (Pehrson and Cohen, 1986), but removal of these cells
does not alter the fertility of the adult urchins (Ransick et
al., 1996). However, intriguing data showing the specific
accumulation in the small micromeres of molecules usually
associated with PGC fate, such as mitochondrial rRNA (Ogawa
et al., 1999) and Vasa protein (C.G.E., unpublished), suggest
that the role of the small micromeres as potential pPGCs
should be re-evaluated.

PGCs in the chordates

The phylum Chordata includes two invertebrate groups,
urochordates (e.g. sea squirts) and cephalochordates (e.g.
Amphioxus), as well as the vertebrates. The origin of the germ
line in urochordates is best understood in solitary ascidians like
Halocynthia roretzi, in which detailed cell lineage studies have
paved the way for contemporary molecular studies (Nishida,
1987; Nishida and Satoh, 1983; Nishida and Satoh, 1985), and
Ciona intestinalis, which has recently joined the ranks of the
‘genomic’ Metazoa (Dehal et al., 2002). Some descriptive
evidence (such as Vasa mRNA and protein localisation,
transcriptional repression of somatic genes) has suggested that
two small blastomeres (B7.6 cells) at the 64-cell stage of C.
intestinalis and H. roretzi are the pPGCs (Fujimura and
Takamura, 2000; Takamura et al., 2002; Tomioka et al., 2002).
An organelle whose ultrastructure resembles that of germ
plasm, called the centrosome-attracting body (CAB), has been
identified in H. roretzi embryos (Isoto and Nishida, 1999;
Nishikata et al., 1999). The CAB is formed in the posterior
vegetal cytoplasm of the two-cell stage embryo, is inherited by
the B7.6 cells during early cleavages, and has been observed to
close in with specific mRNAs (Nakamura et al., 2003).
This observation raises the possibility that somatic and/or germ
cell determinants may be transmitted to putative PGCs via the
CAB. Among the colonial ascidians, it is known that individual
zooids of the colony can exchange germ cells, such that the
PGCs from a single zooid can give rise to almost all of the
trophi of the colony (Stoner and Weissman, 1996).
However, the embryological origin of the PGCs is unknown.

In cephalochordates, the first morphological identification of
germ cells is very late in development in the region of the
gonad anlagen, suggesting that they are epigenetically
determined. Interestingly, an electron-dense region of ooplasm
has been reported to localise to a single blastomere at early
cleavage stages (Holland and Holland, 1992). Further studies
will be necessary to establish whether this blastomere gives rise
to the germ cells, which would provide another example of
germ cell segregation by preformation.

As the evolution of germ cell origin in vertebrates has been
recently reviewed (Johnson et al., 2003), we summarise only
briefly here the general patterns of vertebrate epigenesis and
preformation. Among the Agnatha (see Box 3), lamprey germ
cells are first distinguished at the time of gastrulation, although
their germ layer of origin is uncertain (Beard, 1902a;
Okkelberg, 1921). Few data are available on the embryology
of hagfish, but germ cells in this group have been reported to
arise from the gonadal epithelium (Walvig, 1963). In
cartilaginous fishes, most researchers have first identified germ
cells at late stages of development, and have presumed that
they were of mesodermal origin, although in 1900 John Beard
suggested that their yolk nature meant that they derived from
the blastoderm before mesoderm formation (Beard, 1900;
Beard, 1902b). Extensive studies in zebrafish and some other
teleosts have shown that germ cells form by preformation, but
the examination of other bony fish using a variety of markers
leaves it unclear whether germ cell segregation by preformation
is common to all teleosts, let alone to all ray-
finned fish. Thus it is uncertain whether preformation is the
ancestral mechanism of germ cell formation for all fish.

Very little embryological information is available on
carcin larvae (see Box 3) other than tetrapods, but in
dipnoans (see Box 3), Andrew Johnson and colleagues have
failed to detect a mitochondrial cloud in oocytes of the lungfish
Protoposaurus annectans, which suggests that germ cell
determinants are not localised in this animal before the onset
of embryogenesis. This leads Johnson and colleagues to favour
an epigenetic origin of germ cells in this group (Johnson et al.,
2002).

The living tetrapods include the Amphibia (frogs,
salamanders, newts) and the Amniota (birds, reptiles,
mammals). We have already discussed the descriptive and
The phylogenetic positions of the Chaetognatha and Gnathostomulida are particularly uncertain (dotted lines).

Few amniotes, other than birds and mammals, have been studied in detail. The studies on vasa protein distribution throughout chick development have shown that germ cells are specified during cleavage stages, and are not induced from a subset of epiblast cells around the time of primitive streak formation, as had been thought previously. Most studies on reptiles, including turtles (Fig. 2G), suggest that PGCs in these organisms originate in the extraembryonic endoderm, presumably epigenetically as there is no evidence for a predetermined subset of extraembryonic cells that later differentiate to become PGCs (Fig. 2H) (Risley, 1933). Notwithstanding these data, bird PGCs were also considered to be induced epigenetically from extraembryonic tissue before the convincing quail-chick chimera experiments on uterine stage chick embryos, which showed that chick PGCs were derived from the epiblast well before primitive streak formation (Eyal-Giladi et al., 1981), were published. Subsequent examination of Vasa protein localisation was necessary to establish that chick PGCs are probably specified by preformation. The availability of *vasa* as a molecular marker may allow the origin of PGCs in other sauropsid species (see Box 3) to be clarified. The cross-reacting chicken Vasa antibody developed by Tsunekawa and coworkers (Tsunekawa et al., 2000) also offers exciting possibilities for further study of germ cell origin in reptiles.

The data available on the Eutheria (placental mammals) strongly suggests that the epigenetic segregation of germ cells, which has been so well characterised in mice, is common to all placental mammals. However, embryological studies carried out in the past on the monotremes (see Box 3) and metatherians (see Box 3) have not been able to determine the time or place of germ cell origin in these animals. The possibility that localised determinants may play a role in embryonic pattern formation in marsupials cannot be ruled out (Selwood, 1968), but there is nothing to suggest that these determinants have anything to do with germ cells, whose earliest reported visualisation is at the 12 somite stage with germ cells, whose earliest reported visualisation is at the 12 somite stage using alkaline phosphatase as a marker (Ullmann et al., 1997). The extreme difficulty of obtaining monotreme specimens for study (Caldwell, 1887) means that even modern studies of development in these animals often rely on histological preparations that are ~100 years old (Hughes and Hall, 1998). Overall, it seems likely that, with few or no exceptions, mammals rely upon epigenetic mechanisms to specify germ cells.

**Fig. 3.** Modes of germ cell specification across the Metazoa. Boxes refer to modes of germ cell specification as described in the existing literature: red, epigenesis; blue, preformation; half red, half blue, groups in which some species show preformation and others epigenesis; white, no data. Asterisks indicate phyla in which epigenesis has been claimed, but recent data suggest preformation (see discussion in main text). Phylogeny is modified from Peterson and Eernisse (Peterson and Eernisse, 2001), but many relationships within the Ecdysozoa and Lophotrochozoa remain unresolved by molecular data (Adoutte et al., 2000). The phylogenetic positions of the Chaetognatha and Gnathostomulida are particularly uncertain (dotted lines).
from this summary on the ancestry and evolution of germ cell specification mechanisms.

**Similarities and differences in the germ line across the Metazoa**

In species that segregate PGCs by preformation, the germ line is immortal and continuous from generation to generation, and this makes it tempting to speculate that preformation has a common origin and continuous history. However, closer inspection makes it clear that in only three cases are entire phyla characterised by germ plasm-driven PGC specification (rotifers, nematodes and chaetognaths), and none of these phyla can be considered to be basal to the Metazoa (Fig. 3). Other clades that show PGC segregation via preformation (e.g. dipteran insects, anuran amphibians, archosaurian reptiles) are derived lineages within phyla for which epigenetic specification is likely to be a basal mechanism (Fig. 3).

The data we have reviewed here suggest that PGCs can be segregated at almost any point during embryogenesis: before blastoderm formation; after embryonic rudiment formation but before germ layer separation; after germ layer separation but before gonadogenesis; or after gonadogenesis and continuously throughout adult life. Although many studies are not experimental, and are therefore not conclusive, for most phyla we have been able to combine observations based on the distinctive morphology of germ cells with those based on molecular techniques of PGC identification. In members of 23 out of 28 phyla, PGCs are first observed after embryonic rudiment formation. These observations imply that inductive signals are probably responsible for germ line segregation in these groups. The alternative hypothesis is that, in these groups, a germ line is segregated early, but is not distinguished cytologically, and has not yet been identified. Although this is certainly likely to be true for some groups, there are others where the data argue strongly against it (e.g. nemerteans, holothuroids, acelomorphs). On balance, we believe that epigenesis is likely to be the mode used to segregate germ cells in most animals, including all animals basal to the Bilateria. This suggests that epigenesis is probably the basal mode of germ cell specification for the Metazoa. However, the variability in timing and site of germ cell origin suggests that the specific molecular mechanisms used for inductive signaling are unlikely to be the same in all cases.

**Evolutionary origin of germ cells**

The most obvious similarity of PGCs across phyla is the presence of some kind of aggregate of electron-dense, basophilic bodies in the cytoplasm of germ cells. Such aggregates are widely accepted as markers of germ cells, and in some cases have been shown to confer germ cell fate autonomously on the cells that contain them. These aggregates are variously called dense bodies, nuage, mitochondrial clouds, chromatoid bodies, yolk nuclei or Balbiani bodies, and have been observed at some stage during the development of the germ cells of all phyla examined by electron microscopy (see references in Table S2 at http://dev.biologists.org/supplemental/) (see also Eddy, 1975). The exact relationship between all of these differently named structures has not been determined, but it is possible that they are all different morphological manifestations of the same germ line-specific body. The pluripotent cell types of several basal phyla also contain these dense bodies, and gonia in these phyla are derived from such pluripotent cells.

Several convincing studies have shown that the composition of the electron-dense aggregates found in germ cells is similar in widely divergent phyla. They always contain a combination of RNAs, proteins, endoplasmic reticulum and mitochondria, and may sometimes contain other organelles (such as microtubules) as well. Where studied, the proteins and RNAs localised to these aggregates are products of germ cell-specific genes that are often conserved across divergent phyla (e.g. Bradley et al., 2001). The dynamics of organelle movement during the assembly of these aggregates also shows striking similarity between different animals (Carré et al., 2002; Heasman et al., 1984; Holland and Holland, 1992). Thus primordial germ cells, as a specialised cell type, may well be homologous across all Metazoa, by the criterion that they have retained an ancestral suite of molecular characteristics that define the germ cell lineage.

We suggest that this complex suite of molecular characters, including several gene expression profiles, the subcellular architecture of germ cells and possibly molecular mechanisms of regulating gene activity, is likely to have evolved only once, and thus may constitute a homologous cell identity ‘program’. However, this suite of germ cell characters may be turned on in cells of different germ layer origin, at different times and places during development. This means that neither the mechanisms that trigger germ cell formation, nor the cells in which the ‘program’ is elicited, are homologous.

In bilaterian outgroups and basal Bilateria, the induction of germ cells probably occurred in a population of pluripotent somatic stem cells (similar to the archaeocytes of Porifera, the I cells of Cnidaria and the neoblasts of Acoelomorpha). In higher bilaterian lineages, the same germ cell fate may be elicited at different times and from different cells during development, by a variety of mechanisms. In some derived animal lineages, this mechanism may be maternal segregation of determinants, which include components of the molecular assembly that characterise germ cells. If this view is correct, then we might expect that future investigations on the molecular aspects of germ cell differentiation will continue to reveal conservation of the gene products and cell biological characteristics of germ cells, whereas studies on the mechanisms of PGC segregation in non-model organisms may provide experimental evidence for a diversity of mechanisms that trigger germ cell formation, including epigenetic induction, as well as the segregation of determinants.

The authors thank Ronald Jenner, Max Telford and Mark Martindale for comments on the manuscript, Anne McLaren for the photograph of mouse PGCs, and members of the lab for discussion. C.G.E. thanks the staff of the Woods Hole Marine Biological Laboratory, and of the Balfour and Scientific Periodicals Libraries of the University of Cambridge for their help in obtaining the older literature. Part of the work was done at the Marine Biological Laboratory in Woods Hole, with the help of a Carl Zeiss post-course research fellowship awarded to C.G.E. C.G.E. is funded by the Wellcome Trust.

**References**

Heymons, R. (1901). Entwicklungsgeschichte der Scolopender. Zoologica 33, 1-244.
Heymons, R. (1901). Entwicklungsgeschichte der Scolopender. Zoologica 33, 1-244.
Heymons, R. (1901). Entwicklungsgeschichte der Scolopender. Zoologica 33, 1-244.
Heymons, R. (1901). Entwicklungsgeschichte der Scolopender. Zoologica 33, 1-244.


