

The notion of the Cambrian pananimalia genome

SUSUMU OHNO

Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, CA 91010-3011

Contributed by Susumu Ohno, May 1, 1996

ABSTRACT The toil by photosynthesizing cyanobacteria and blue-green algae of nearly three billion years appeared to have finally resulted in the sufficient accumulation of molecular oxygen. So, the stage was set for the emergence, at the ocean bottom, of diverse animals that were consumers of molecular oxygen. It now appears that this Cambrian explosion, during which nearly all the extant animal phyla have emerged, was of an astonishingly short duration, lasting only 6–10 million years. Inasmuch as only a 1% DNA base sequence change is expected in 10 million years under the standard spontaneous mutation rate, I propose that all those diverse animals of the early Cambrian period, some 550 million years ago, were endowed with nearly identical genomes, with differential usage of the same set of genes accounting for the extreme diversities of body forms. Some of the more pertinent genes that are thought to be included in the Cambrian pananimalia genome are as follows. (i) A gene for lysyloxidase that, in the presence of molecular oxygen, crosslinked collagen triple helices to produce ligaments and tendons, thus contributing to the stout bodies of the Cambrian animals. (ii) Genes for hemoglobin; these internal transporters of molecular oxygen are today seen sporadically in members of diverse animal phyla. (iii) The *Pax-6* gene for eye formation; the eyes of a ribbon worm to a human are organized by this gene. In animals without eyes, the same gene organizes other sensory systems and organs. (iv) A series of *Hox* genes for the anterior–posterior (cranio–caudal) body plans: these genes are also present in all phyla of the kingdom Animalia.

Considering the age of our solar system, which is said to be 4.7 billion years, it is rather remarkable that abundant microbial fossils are already evident in sedimentary and volcanic rocks that are between 3.0 to 3.7 billion years old (1). Furthermore, these earliest microbes need not have been all prokaryotic, since certain blue-green algae have been shown to be as small as bacteria; e.g., *Ostreococcus tauri*, containing one mitochondrion and one plastid each, measured less than 1 μm in length (2). It is worth noting that the cell size of several 3.46 billion-year-old filamentous microbes such as *Primaevifilum amoenum* exceeded 5 μm in length. Yet, they were thought to be prokaryotic cyanobacteria rather than eukaryotic blue-green algae (1). At any rate, it would be safe to assume that the release of molecular oxygen, at least by photosynthesizing cyanobacteria and possibly by blue-green algae, commenced about 3.5 billion years ago.

Yet, it appears that to set the stage for the emergence of animals, the consumer of molecular oxygen, at the ocean bottom, the continuous toil of nearly 3 billion years by these photosynthesizing organisms were required. Once the stage was set at about 550 million years ago, nearly every extant phylum of the kingdom Animalia emerged within the time span of 6–10 million years (3). Thus, the term Cambrian explosion became more real than originally intended. As a prelude to the Cambrian explosion, jellyfish-like creatures and sessile colonial Cnidaria that left strikingly leaf-like fossils of

the phylum Coelenterata and of *Dickinsonia*, a radially segmented flat stone-like worm with an affinity toward the phylum Annelida appeared in the Ediacarian period that immediately preceded the Cambrian (4). Sponges of the phylum Porifera might have made even earlier appearances, but there appears to have been a major extinction at the Ediacarian–Cambrian boundary. Accordingly, only a few Ediacarian survivors were represented in the early Cambrian fauna (5). In the early Cambrian fauna, aside from various worms of the phyla Nemertina, Platyhelminthes, Annelida, etc., there were members of the phylum Echinodermata, whereas the phylum Mollusca was represented by shelled bivalves as well as by squid-like cephalopods. The phylum Arthropoda contained a large variety of crustacean-like creatures. More remarkable were *Opabinia regalis*, with five eyes, and the 30-cm long *Anomalocaris*, which is said to have been the largest predator of the early Cambrian (5, 6). Of the three tiers of the phylum Chordata, the early Cambrian fauna were originally thought to contain tunicate-like urochordates only. However, 525 million-year-old *Chengjiang* fauna in the Yunnan Province of China readily yielded fossils of *Yunnanozoon lividum*, an amphioxus-like cephalochordate (7). The possessor of a set of comb-like teeth was defined as conodonts, and they were found among the early Cambrian fauna. Recently, this conodont was identified as a large-eyed, eel-like fish, ≈ 20 cm long, representing the most primitive jawless stage of vertebrate evolution (8). Although *Promissum pulchrum* itself was of the Ordovician origin, a similar agnathan fish must have necessarily been present in the early Cambrian period (8).

Reasons for Invoking the Presence of the Cambrian Pananimalia Genome. Assuming the spontaneous mutation rate to be generous 10^{-9} per base pair per year and also assuming no negative interference by natural selection, it still takes 10 million years to undergo 1% change in DNA base sequences. It follows that 6–10 million years in the evolutionary time scale is but a blink of an eye. The Cambrian explosion denoting the almost simultaneous emergence of nearly all the extant phyla of the kingdom Animalia within the time span of 6–10 million years can't possibly be explained by mutational divergence of individual gene functions. Rather, it is more likely that all the animals involved in the Cambrian explosion were endowed with nearly the identical genome, with enormous morphological diversities displayed by multitudes of animal phyla being due to differential usages of the identical set of genes. This is the very reason for my proposal of the Cambrian pananimalia genome. This genome must have necessarily been related to those of Ediacarian predecessors, representing the phyla Porifera and Coelenterata, and possibly Annelida. Being related to the genome possessed by the first set of multicellular organisms to emerge on this earth, it had to be rather modest in size. It should be recalled that the genome of modern day tunicates, representing subphylum Urochordata, is made of 1.8×10^8 DNA base pairs, which amounts to only 6% of the mammalian genome (9).

The following are the more pertinent of the genes that were certain to have been included in the Cambrian pananimalia genome.

Lysyloxidase for the Stout Cambrian Bodies via Collagen Crosslinking. As already noted, molecular oxygen was the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

likely key to the Cambrian explosion, and indeed there was, and still is, one key enzyme that requires molecular oxygen. Compared with their Ediacarian predecessor, Cambrian animals in general were characterized by their much stouter bodies (5). The stoutness of the body is likely due to the formation of ligaments and tendons, which in turn requires the crosslinking of collagen triple helices. As schematically illustrated in Fig. 1, the initial stage in the formation of intermolecular bonds is the oxidative deamination of specific lysine or hydroxylysine residues. The aldehydes, thus produced, form the aldimine bound with ϵ -amino groups of hydroxylysine and, by so doing, complete the crosslinking of collagen triple helices. Molecular oxygen and copper-dependent lysyloxidase are required for the bond formation noted above, and the importance of lysyloxidase is most dramatically illustrated by human patients with Ehler-Danlos syndrome type VI. In the functional absence of lysyloxidase, skin, ligaments, and tendons lose all their tensile strength (10). It would thus appear that the lysyloxidase gene was one of the key genes in the pananimalia genome that made the Cambrian explosion possible.

Hemoglobins as Transporters of Molecular Oxygen. By displacing a pair of cysteine residues that used to hold heme with a pair of histidines, hemoglobin molecules that readily bound and then released molecular oxygen in peripheral tissues were born. However, since Cambrian animals tended to be of small body sizes, the need for internal transporters of molecular oxygen might not have been so universal. Further complication was introduced by an apparently simultaneous invention of an alternative to hemoglobins in the form of hemocyanins, copper-containing respiratory pigments. Nevertheless, the fact is that possessors of hemoglobins are distributed over diverse animal phyla (11). For example, earthworms of the class Oligochaeta and the phylum Annelida are uniformly red-blooded because of hemoglobins dissolved in their circulating fluid. It should be recalled that this phylum might have already been represented in the Ediacarian period in the form of *Dickinsonia* (4). Of dipteran insects, no hemoglobin can be found either in larvae or adults of the genus *Drosophila*, whereas aquatic larvae of its close relative *Chironomus* are

richly endowed with a variety of hemoglobin molecules. Crustaceans of this phylum, on the other hand, tend to utilize hemocyanins as respiratory pigments.

Of the phylum Mollusca, cephalopods tend also to employ hemocyanins, whereas the pinkness of certain bivalves is due to the presence of hemoglobin molecules. Of the three tiers of the phylum Chordata, no trace of hemoglobin is found either in urochordates or in cephalochordates. On the contrary, in all vertebrates from fish to mammals, hemoglobins are encased in a very specialized cell type, erythrocytes. This, however, is not an invention unique to vertebrates, for certain sea cucumbers of the class Holothuroidea of the phylum Echinodermata have nucleated erythrocytes containing hemoglobins (12). All in all, there remains little doubt that a gene or genes encoding hemoglobins were contained in the Cambrian pananimalia genome, although many small-body-sized animals chose not to utilize this respiratory pigment.

A Set of Genes for Glass (Silicified) Skeleton. Cambrian bivalves of the phylum Mollusca utilized calcium carbonate for their shells (exoskeleton), whereas endoskeletons of the Cambrian vertebrate in the form of agnathan fish likely contained calcium phosphate. Yet, there is a third material that has been used for the construction of skeletons. Diatomites are silicified skeletal remains of minute, unicellular or colonial algae belonging to the family Bacillariophyceae. Glass sponges belong to the class Hexactinellida of the phylum Porifera. As indicated by their name, their skeletons are silicified. Thus, a link between algae and animals is established. Quite recently, chain mail nematodes of the order Desmoscolecoida were found to be equipped with a series of roughly 18 silicified endoskeletal rings (13). The occurrence of glass (silicified) skeletons in two diverse phyla implies that further search would likely reveal widespread use of glass skeletons within the kingdom Animalia.

Pax-6 Gene for the Formation of Eyes and Other Sensory Organs. As with possessors of hemoglobins, possessors of eyes are distributed sporadically over diverse animal phyla. For example, mostly marine ribbon worms of the phylum Nemeritina have a pair of eyes. Of various flatworms of the phylum Platyhelminthes, planaria of the order Tricladida, the class

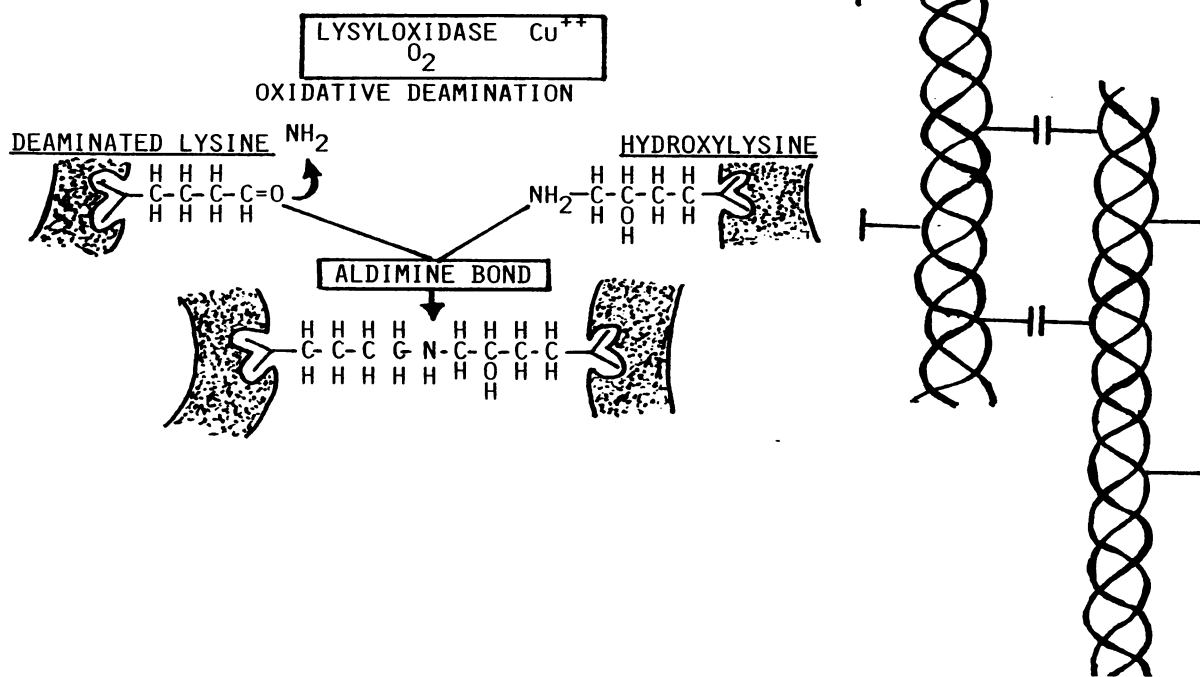
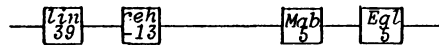


FIG. 1. A schematically illustrated crosslinking of collagen triple helices by copper-dependent lysyloxidase in the presence of molecular oxygen (10).

Phylum: Nemathelminthes

Class: Nematoda

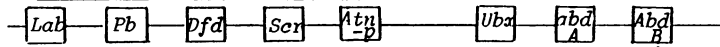
Caenorhabditis elegans



Phylum: Arthropoda

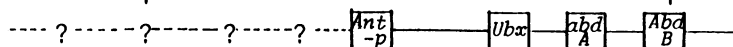
Class: Insecta

Drosophila melanogaster



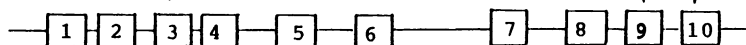
Class: Crustacea

Artemia franciscana



Phylum: Chordata Subphylum: Cephalochordata

Amphioxus species



Subphylum: Vertebrata

Class: Mammalia

Mus musculus

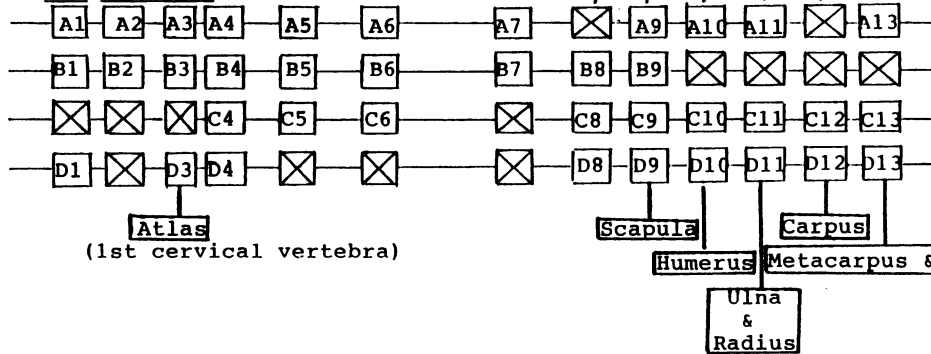


FIG. 2. Numbers and types of Hox genes present in five divergent species belonging to different phyla, subphyla, and classes are schematically identified. Corresponding Hox genes of different animals are, as a rule, aligned on the same vertical line. Further kinships among individual Hox genes of divergent animals are indicated by trans-specific connecting lines. In the case of the subphylum Vertebrata, represented by the domesticated mouse *Mus musculus*, defunct Hox genes are identified by crosses, and somite derivatives controlled by individual groups of Hox genes, where known, are identified (e.g., *Hox-a3* and *Hox-d3* together control the development of atlas, the first cervical vertebra, whereas the *Hox-b3* gene apparently became redundant; refs. 22-27).

Turbellaria are endowed with a pair of eyes that readily regenerate. Among members of the phylum Mollusca, cephalopods and gastropods have eyes, whereas eyes are not present among bivalves. With regard to the phylum Arthropoda, only adults of the class Insecta are endowed with a pair of compound eyes, whereas their larvae tend to have odd numbers of simple eyes on each side; somewhat reminiscent of *O. regalis* of the early Cambrian with five eyes (5). In view of the above, it is all the more remarkable that the eyes of all animals are formed under the direction of the same gene and that this gene persists even in animals without eyes, controlling the development of other sensory systems and organs (14, 15). Pax-6 protein belongs to a particular family of DNA binding proteins, and as far as their paired domains and homeodomains are concerned, 94% amino acid sequence identity is maintained between this protein of *Drosophila melanogaster*, on one hand, and those of humans, the mouse and the quail on the other (14). Fruit flies, made homozygous deficient for this gene, become eyeless, whereas rats and mice, heterozygous for a deficiency of this gene, form small eyes. In the case of human beings heterozygously deficient for this gene, irises fail to develop, hence aniridia (14, 15). When rats were made ho-

mozygous for *Pax-6* deficiency, they became eyeless as well as noseless (15). It follows that the *Pax-6* gene, in addition to controlling the formation of eyes in all animals, may also control the development of other sensory systems and organs. Indeed, the *Pax-6* gene was found in all the animal phyla studied. With regard to those with eyes, a ribbon worm (*Lineus sanguineus*) of the phylum Nemertina, as well as planaria of the phylum Platyhelminthes expressed their *Pax-6* gene locally during eye formation (16). With regard to those without eyes, a nematode (*Caenorhabditis elegans*) of the phylum Nemathelminthes and a sea urchin of the phylum Echinodermata possessed *Pax-6* genes (17-19). In the case of the former, there were two *Pax-6* genes (17, 18). There remains little doubt that either one or a series of *Pax-6* genes were included in the Cambrian pananimalia genome.

A Set or Sets of Genes for the Anterior-Posterior (Cranio-Caudal) Body Plan. The *Antp* (*Antenapedia*) series of five genes (*Antp*, *Scr*, *Dfd*, *Pb*, and *Lab*) and the *Ubx* (*Bithorax*) series of three genes (*Ubx*, *Abd-A*, and *Abd-B*) were originally found as developmental controllers of body segments of *D. melanogaster* (20). In fact, arthropodes of the Cambrian explosion were predominantly crustacean-like creatures (5, 6). It

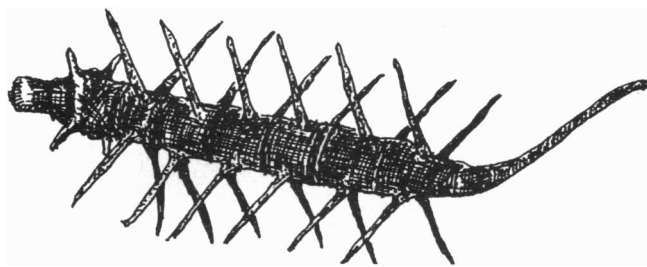


FIG. 3. One of the most odd-looking early Cambrian animals is *H. sparsa*, whose body was made of six or seven identical body segments; each segment sported four identical appendages, thus creating levo-dextro as well as dorso-ventral perfect symmetry (5). The genome of this animal might have contained either one *Hox* gene or a set of eight *Hox* genes, as in arthropods, with each of the eight genes being identical to each other.

would follow that the most ancestral body plan of the phylum Arthropoda is to be found among the crustaceans. The body plan of the brine shrimp (*Artemia franciscana*) of the subclass Brachiopoda was found to utilize at least four different *Hox* genes that are equivalent of drosophilean *Antp*, *Ubx*, *Abd-A*, and *Abd-B* (21). It is of interest to note that *Hox* genes are found in the genome of animals whose body are not at all segmented, e.g., *C. elegans* (nematoda) of the phylum Nematelminthes. However, it should be recalled that chain mail nematodes of the order Desmoscolecoida were equipped with a set of roughly 18 glass skeletal rings (13). Four *Hox* genes, identified as *lin-39*, *ceh-13*, *mab-5*, and *egl-5* were found in the genome of the above-named nematoda species (22). It should be recalled that *Pax-6* genes were also present in this species without eye development. In the case of this nematoda species, functional conservation over 500 million years of individual *Hox* genes was also revealed, for it was shown that defects manifested by affected nematodes due to defunct *lin-36* and *mab-5* genes can be remedied by transfection with either of the two drosophilean *Hox* genes, *Scr* and *Antp*, respectively (22). Similarly, genetic defects due to defunct *Dfd* of *D. melanogaster* were remedied by transfection with the human *HOX-b4* gene (23).

As to the phylum Chordata, one amphioxus species representing the subphylum Cephalochordata was shown to have a set of 10 *Hox* genes, two more than the fruit fly as shown in Fig. 2 (24). In sharp contrast to all other animal phyla studied so far, members of the subphylum Vertebrata were characterized by having two or more sets of *Hox* genes, with teleost fish to mammals uniformly having four sets of *Hox* genes. Inasmuch as large chromosomal regions surrounding these four sets of *Hox* genes of humans and the mouse are homeologous (25), it would appear that these four sets of *Hox* genes came into being by two successive tetraploidization events that occurred in Devonian ancestors of the vertebrates (9).

The original number of *Hox* genes in each of the four vertebrate sets appeared to have been 13. Inasmuch as the maximal number of somite pairs finally reached by vertebrate embryos at about the limb bud stage approaches 50, it is an interesting coincidence that four sets of 13 *Hox* genes each add up to a total of 52 *Hox* genes. Yet, the opportunity to control each of the 50 odd pairs of the somites with an individual *Hox* gene had apparently been forfeited. Both in humans and in mouse, 27% of the potential 52 *Hox* genes had already become defunct, thus joining the rank of pseudo genes (Fig. 2). In fact, the remaining *Hox* genes of mammals appear to control, in a redundant manner, only the fate of the first dozen or so pairs of somites initially formed at the neural groove stage. In the mouse, two of the three surviving *Hox-3* genes cooperatively

control development of atlas, the first cervical vertebra (26). The last *Hox* series, *Hox-13*, is also represented by three survivors, and these three *Hox-13* genes control development of metacarpal and phalangeal bones of the forelimb (27). As illustrated in Fig. 2, four sets of *Hox* genes had largely been wasted by vertebrates, but this waste might be viewed as manifestation of the historical burden. *Hox* genes initially included in the Cambrian pananimalia genome were designed to deal with only a dozen or so body segments.

What if a Cambrian Animal Either Had a Single *Hox* Gene or a Set of Identical *Hox* Genes? Inasmuch as one *Hox* gene should suffice for the body plan of animals whose bodies are made exclusively of successive, but identical, body segments (i.e., millipedes with two pairs of legs in each body segment), Lewis (20) thought of millipede-like creatures as the ancestral form of the phylum Arthropoda. In view of the possession of *Hox* genes by all the animal phyla, the Lewis's view (20) should be enlarged to include the entire animal kingdom. Was there a Cambrian animal with this ancestral body plan? The one that immediately comes to my mind, at least, is *Hallucigenia sparsa*, whose body is made of six or seven identical body segments, each of which sport ventral as well as dorsal pairs of identical stilt-like appendages, thus maintaining levo-dextro as well as dorso-ventral symmetry (5). Shown in Fig. 3 is my version of *H. sparsa*.

- Schopf, J. W. (1993) *Science* **260**, 640–646.
- Courties, C., Vaquer, A., Troussellier, M., Lautier, J., Chretiennot-Dimet, M. J., Neveux, J., Machado, C. & Claustre, H. (1994) *Nature (London)* **370**, 255.
- Gould, S. J. (1995) *Nature (London)* **377**, 681–682.
- Cloud, P. & Glaessner, M. F. (1982) *Science* **218**, 783–792.
- Conway Morris, S. (1989) *Science* **246**, 339–346.
- Briggs, D. (1994) *Science* **264**, 1283–1284.
- Chen, J.-Y., Dzik, J., Edgcombe, G. D., Ramskold, L. & Zhou, G.-Q. (1995) *Nature (London)* **377**, 720–722.
- Gabbot, S. E., Aldridge, R. J. & Theron, J. M. (1995) *Nature (London)* **374**, 800–803.
- Ohno, S. (1970) *Evolution by Gene Duplication* (Springer, New York).
- Bailey, A. J., Robins, B. P. & Balian, G. (1974) *Nature (London)* **251**, 105–109.
- Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 4.
- Ohue, U. (1936) *Sci. Rep. Tohoku Imper. Univ. Biol.* **2**, 223–239.
- Shiroyama, Y. (1996) *Kagaku (Tokyo)* **66**, 312–317.
- Quiring, R., Waldorf, U., Kloter, U. & Gehring, W. (1994) *Science* **265**, 785–787.
- Matsuo, T., Osumi-Yamashita, N., Noji, S., Ohuchi, H., Koyama, E., Myokai, F., Matsuo, N., Toniguchi, S., Doi, H., Iseki, S., Ninomiya, Y., Fujiwara, M., Eatanabe, T. & Eto, K. (1993) *Nat. Genet.* **3**, 299–304.
- Loosli, F., Kumita-Cunisse, M. & Gehring, W. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 2658–2663.
- Chrisholm, A. D. & Horovitz, H. R. (1995) *Nature (London)* **377**, 52–55.
- Zhang, Y. & Emmons, S. W. (1995) *Nature (London)* **377**, 55–59.
- Czerny, T. & Busslinger, M. (1995) *Mol. Cell. Biol.* **15**, 2858–2871.
- Lewis, E. B. (1978) *Nature (London)* **276**, 567–570.
- Averof, M. & Akam, M. (1995) *Nature (London)* **376**, 420–423.
- Hunter, C. P. & Kenyon, C. (1995) *Nature (London)* **377**, 229–232.
- McGinnis, N., Kuziora, M. A. & McGinnis, W. (1990) *Cell* **63**, 969–976.
- Garcia-Fernandez, J. & Holland, P. W. H. (1994) *Nature (London)* **370**, 563–566.
- Lundin, L. G. (1993) *Genomics* **16**, 1–19.
- Condie, B. G. & Cepecchi, M. R. (1994) *Nature (London)* **370**, 304–307.
- Davis, A. P., Witte, D. P., Hsieh-Li, H. M., Patter, S. S. & Capecchi, M. R. (1995) *Nature (London)* **375**, 791–795.