Commentary

Functional gene cassettes in development

Yuh Nung Jan and Lily Yeh Jan

Departments of Physiology and Biochemistry and Howard Hughes Medical Institute, University of California, San Francisco, CA 94143-0724

It has been noted in recent years that groups of genes are often used at multiple stages and multiple places of development. Comparison of the functions of each group of genes in different developmental context gave rise to the notion of "functional gene cassettes." These groups of genes can often be viewed as sets of tools; each group is suited to carry out a certain operation and may serve similar functions in several developmental processes. For example, neurogenic genes can be thought of as a set of tools for singling out cells from a group of equivalent cells. The affected cells could be neuroectodermal cells, myoblasts, or a number of other cell types encountered at various stages of development (see below).

A corollary is that if one notices that one member of a group is used in any developmental process, then there is good reason to suspect that other members may also be used. This simple notion provides a rather useful way of attempting to identify genes that control a hitherto poorly characterized developmental process.

In the following, we will illustrate this notion with examples of three groups of genes. All three groups have important functions in controlling early steps of *Drosophila* neurogenesis. Each group is also utilized in other developmental processes.

A Group of Helix–Loop–Helix (HLH) Proteins Controls the Initial Steps of Sex Determination and Neurogenesis

On the surface, neurogenesis and sex determination appear to be rather different biological processes. Therefore it was initially a surprise that a gene, daughterless (*da*), which had been known previously to be required for proper sex determination (1), was found to be essential for the formation of the entire *Drosophila* peripheral nervous system (PNS) (2). Recent studies have revealed extensive overlap between the genetic control of sex determination and that of neural development (3–6); a group of HLH proteins has been found to regulate the initial steps of both processes.

In *Drosophila* neurogenesis, the founding event is the expression of proneural genes (reviewed in refs. 7–9). Proneural genes are expressed in clusters of ectodermal cells and thereby endow those cells with the potential of developing into neuronal precursors. In other words, where proneural genes are expressed determines where the nervous system develops. All proneural genes that are required for PNS development (achaete, scute, atonal) encode basic–HLH (bHLH) proteins. The proneural genes function with a partner, daughterless, which encodes a ubiquitously distributed bHLH protein. They form heterodimers and control the expression of another group of genes. The latter are called neuronal precursor genes because they are expressed in neuronal precursors and are thought to control the neuronal differentiation of neuronal precursors and their progenies. The proneural genes in turn are regulated by a number of genes, including two negative regulators, hairy (*h*) and extramacrochaete (*emc*) (10), which encode HLH proteins. Unlike the proneural gene products, emc protein lacks basic domain. It appears that emc protein negatively regulates proneural genes by forming heterodimers with achaete (ac) or daughterless (da) protein and thereby sequestering the functional da/ac heterodimers (11–13).

In *Drosophila*, sex is determined by the X chromosome (X) to autosome (A) ratio (X/A ratio) (14). In the sex determination pathway (reviewed in refs. 15 and 16), the key regulatory gene is Sex lethal (*Sxl*). It controls three sets of genes, which in turn control somatic sex determination, dosage compensation, and germ-line sex determination. Whether *Sxl* is on or off is determined by the X/A ratio. In females, there are two X chromosomes and two sets of autosomes. The X/A ratio is one and *Sxl* is turned on. This leads to female development. In males, there is only one X chromosome. The X/A ratio is 0.5 and *Sxl* is off. This leads to male development. What measures X/A ratio? Several X-linked genes, including sisterless-a (*sis-a*) and sisterless-b (*sis-b*), have been found to be numerator elements (17), whereas an autosomal gene, deadpan (*dpn*), functions as a denominator element. Those "counting elements" serve to measure X/A ratio and thereby control the initial expression of *Sxl* at the transcriptional level. It is those regulators of

Sxl that are extensively shared between the sex determination and neurogenesis pathways. *sis-b* is scute; one of the proneural genes (3–5), deadpan, is also a neuronal precursor gene, which has strong sequence homology with hairy (18).

The comparison of sex determination and neurogenesis is summarized in Fig. 1. In sex determination, the first gene in the pathway that appears to be specific for sex determination pathway is *Sxl*. In the neurogenesis pathway, the group of neuronal precursor genes can be viewed as the corresponding genes; each regulates some aspect of neuronal differentiation. In sex determination, numerator elements are positive regulators of *Sxl*. The proneural genes are the corresponding positive regulators of neuronal precursor genes. *sc* is a shared element; so is *da*, which is a positive cofactor. Both pathways also share negative regulators *emc* and *h* (and its close relative *dpn*).

It may be evolutionary opportunism that a group of HLH proteins that is well suited to serve as a sensitive bistable genetic switch is used to control the transcription of key regulatory genes. The HLH proteins control *Sxl* early in development (between the first and second hour of embryogenesis) for sex determination and they control neuronal formation and differentiation later in development (approximately between the third and sixth hour during embryogenesis).

Neurogenic Genes Function to Single Out Cells from an Equivalence Group

The cells within a cluster that express a proneural gene (called a proneural cluster) can be thought of as cells of an equivalence group. Within a proneural cluster, the cells compete with each other such that only a subset of cells is singled out to develop into neuronal precursors. This singling out process is mediated by cell–cell interaction interpreted through the action of neurogenic genes. Loss of function mutation of any of the neurogenic genes, Notch (*N*), Delta (*Dl*), Enhancer of Split [*E(spl)*], neuralized (*neu*), mastermind (*mam*), and big brain (*bib*), causes a greater number of the cells in the proneural clusters to take the neuronal precursor fate (reviewed in refs. 9, 19–21).
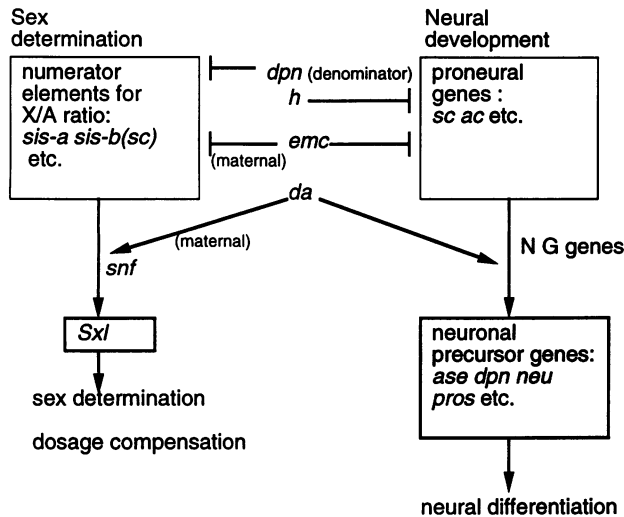
FIG. 1.   Comparison of sex determination and neural development pathways (from ref. 6). Four genes are shared by the two pathways: *da*, *sc/sis-b*, *dpn*, and *emc*. NG, neurogenic; *snf*, sans fille; *ase*, asense; *neu*, neuralized; *pros*, prospero.

Neurogenic genes are so named because they were initially identified by their neuronal phenotype (22). However, recent studies have revealed that the functions of neurogenic genes are not restricted to the nervous system. In fact, in almost every tissue examined in all three germ layers (ectoderm, mesoderm, and endoderm), neurogenic genes are involved in their development (23–27). In most cases examined, the function of neurogenic genes can be interpreted as a requirement to single out a subset of cells from an equivalence group of cells for a particular fate. Thus, in neuroectoderm, neurogenic genes are required to single out cells from within proneural clusters to form neuronal precursors, leaving the remaining cells of proneural clusters to develop into epidermal cells. In myogenesis, neurogenic genes function to single out particular types of muscle precursor cells, such as the ones that express nautilus (*nau*) genes. In loss of function neurogenic mutants, extra *nau*$^+$ cells are formed, presumably at the expense of other types of muscle precursor cells (23). In oogenesis, *N* and *Dl* (and possibly other neurogenic genes) are involved in singling out a subset of follicle cells to become polar cells. In *N*$^-$ and *Dl*$^-$ mutants, extra polar cells are formed, presumably at the expense of other types of follicle cells (24).

It should be noted, however, that the function of neurogenic genes may be even broader than singling out cells from an equivalent group. In certain ectodermally and endodermally derived tissues, neurogenic genes function to mediate cell–cell interaction needed to acquire or to maintain an epithelial phenotype (27). There is also evidence suggesting that *N* and *Dl* are involved in axon guidance by mediating the cell–cell interaction be-

tween axons and the substrate cells on which axons grow (28).

Rhomboid Provides a Spatial Cue in Cell Induction Via the Epidermal Growth Factor (EGF) Receptor Signaling Pathway

The two examples described above led to the notion of a functional gene cassette. The third example illustrates the usefulness of this notion in the identification of genes that control a developmental process that has not yet been well characterized. The biological problem in question is how axes are set up during oogenesis. How anterior–posterior (A–P) and dorsal–ventral (D–V) axes are established during *Drosophila* embryogenesis has been fairly well understood (reviewed in ref. 29). This pushes the problem back in time—i.e., the question becomes how A–P and D–V axes are set up prior to embryogenesis (i.e., during oogenesis).

For the D–V axis, Schüpbach (30) made the important observations that D–V patterning during oogenesis depends on the transfer of spatial information between the germ-line cells and the somatic follicle cells and that EGF receptors are involved. In *Drosophila*, the EGF receptor signaling pathway involves a group of genes called the spitz group, including spitz, rhomboid, pointed, and Star (Fig. 2; see also refs. 31 and 32). In this pathway, spitz protein is probably a ligand because it shows sequence similarity with EGF (33). rhomboid encodes a membrane protein that probably functions as a cofactor of EGF receptor (34). Mutants of the spitz group share various phenotypes in several tissues, including ventral cuticle, PNS, muscle pattern, and midline glia (33–36). Those phenotypes

are probably due to defects in inductive signaling between neighboring cells.

In embryos, the spatial cue of the EGF receptor signaling pathway appears to be provided by rhomboid. The phenotypes of the spitz group of genes are restricted to a few places in the embryo [e.g., ventral midline, lateral regions of PNS (33–36)]. This spatial restriction coincides with the rhomboid expression pattern (34), whereas the EGF receptor and the putative ligand (spitz protein) are distributed fairly ubiquitously (33, 37). Since EGF receptor has been shown to be involved in D–V patterning during oogenesis (30), one might apply the notion of functional gene cassette and predict that the entire spitz group of genes is involved in this process. Further, since rhomboid is the member of the group that provides the spatial cue during embryogenesis, one may expect it to serve the same function during oogenesis. This reasoning prompted Ruohola-Baker *et al.* (31) to examine the possibility that rhomboid provides a spatial cue in setting up D–V axes during oogenesis. Indeed, rhomboid protein is localized to the dorsal–anterior subset of follicle cells surrounding the oocyte. Loss of rhomboid function causes ventralization of the egg chambers, whereas ectopic expression leads to dorsalization. Thus, rhomboid appears to provide a spatial cue in D–V patterning during oogenesis.

Concluding Remarks

Pleiotropy of a gene function used to be somewhat dreaded by developmental biologists because this puts a greater burden on the investigator to find out whether a mutant phenotype is the primary result of the mutation or a consequence several steps removed from the primary cause. However, with increasingly more sophisticated methods at the investigator's disposal to control the spatial and temporal expression pattern of normal and mutated genes (38, 39), this has become less of a problem. It now appears to be more of a rule rather than an exception that a group of genes is used in several different developmental contexts. As illustrated in the examples in this article, the investigators can actually
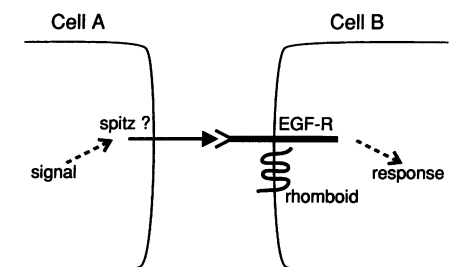

FIG. 2.   EGF receptor signaling pathway mediates induction between neighboring cells.

Commentary: Jan and Jan

Proc. Natl. Acad. Sci. USA 90 (1993)   8307

turn the situation to their advantage in transferring knowledge gained from studying one biological process to another.

1. Cline, T. W. (1976) *Genetics* **84**, 723–742.
2. Caudy, M., Grell, E. H., Dambly-Chaudiere, C., Ghysen, A., Jan, L. Y. & Jan, Y. N. (1988) *Genes Dev.* **2**, 843–852.
3. Torres, M. & Sanchez, L. (1989) *EMBO J.* **8**, 3079–3086.
4. Parkhurst, S. M., Bopp, D. & Ish-Horowicz, D. (1990) *Cell* **63**, 1179–1191.
5. Erickson, J. W. & Cline, T. W. (1991) *Science* **251**, 1071–1074.
6. Younger-Shepherd, S., Vaessin, H., Bier, E., Jan, L. Y. & Jan, Y. N. (1992) *Cell* **70**, 911–922.
7. Cabrera, C. V. (1992) *Development* **115**, 893–901.
8. Campuzano, S. & Modolell, J. (1992) *Trends Genet.* **8**, 202–208.
9. Ghysen, A., Dambly-Chaudière, C., Jan, L. Y. & Jan, Y. N. (1993) *Genes Dev.* **7**, 723–733.
10. Botas, J., Moscoso del Prado, J. & Garcia-Bellido, A. (1982) *EMBO J.* **1**, 307–310.
11. Cabrera, C. V. & Alonso, M. C. (1991) *EMBO J.* **10**, 2965–2973.
12. Van Doren, M., Ellis, H. M. & Posakony, J. W. (1991) *Development* **113**, 245–255.
13. Cubas, P. & Modolell, J. (1992) *EMBO J.* **11**, 3385–3393.
14. Bridges, C. B. (1921) *Science* **54**, 252–254.
15. Baker, B. S. (1989) *Nature (London)* **340**, 521–524.
16. Steinmann-Zwicky, M., Amrein, H. & Nöthiger, R. (1990) *Adv. Genet.* **27**, 189–237.
17. Cline, T. W. (1988) *Genetics* **119**, 829–862.
18. Bier, E., Vaessin, H., Younger-Shepherd, S., Jan, L. Y. & Jan, Y. N. (1992) *Genes Dev.* **6**, 2137–2151.
19. Campos-Ortega, J. A. (1988) *Trends Neurosci.* **11**, 400–405.
20. Simpson, P. (1990) *Development* **109**, 509–519.
21. Artavanis-Tsakonas, S. & Simpson, P. (1991) *Trends Genet.* **7**, 403–408.
22. Lehmann, R., Jimenez, F., Dietrich, U. & Campos-Ortega, J. A. (1983) *Roux's Arch. Dev. Biol.* **192**, 62–74.
23. Corbin, V., Michelson, A. M., Abmayr, S. M., Neel, B., Alcamo, E., Maniatis, T. & Young, M. W. (1991) *Cell* **67**, 311–323.
24. Ruohola, H., Bremer, K. A., Baker, D., Swedlow, J. R., Jan, L. Y. & Jan, Y. N. (1991) *Cell* **66**, 433–449.
25. Xu, T., Caron, L. A., Fehon, R. G. & Artavanis-Tsakonas, S. (1992) *Development* **115**, 913–922.
26. Bender, L. B., Kooh, P. J. & Muskavitch, M. A. T. (1993) *Genetics* **133**, 967–978.
27. Hartenstein, A. Y., Rugendorff, A., Tepass, U. & Hartenstein, V. (1992) *Development* **116**, 1203–1220.
28. Giniger, E., Jan, L. Y. & Jan, Y. N. (1993) *Development* **117**, 431–440.
29. St. Johnston, D. & Nüsslein-Volhard, C. (1992) *Cell* **68**, 201–219.
30. Schüpbach, T. (1987) *Cell* **49**, 699–707.
31. Ruohola-Baker, H., Grell, E., Chou, T.-B., Baker, D., Jan, L. Y. & Jan, Y. N. (1993) *Cell* **73**, 953–965.
32. Sturtevant, M. A., Roark, M. & Bier, E. (1993) *Genes Dev.* **7**, 961–973.
33. Rutledge, B. J., Zhang, K., Bier, E., Jan, Y. N. & Perrimon, N. (1992) *Genes Dev.* **6**, 1503–1517.
34. Bier, E., Jan, L. Y. & Jan, Y. N. (1990) *Genes Dev.* **4**, 190–203.
35. Mayer, U. & Nüsslein-Volhard, C. (1988) *Genes Dev.* **2**, 1496–1511.
36. Klämbt, C., Jacobs, J. R. & Goodman, C. S. (1991) *Cell* **64**, 801–815.
37. Zak, N. B., Wides, R. J., Schejter, E. D., Raz, E. & Shilo, B. (1990) *Development* **109**, 865–874.
38. Golic, K. G. & Linquist, S. (1989) *Cell* **59**, 499–509.
39. Brand, A. & Perrimon, N. (1993) *Development* **118**, 401–415.