9. J. K. Mitchell *et al.*, in *Apollo 17 Preliminary Science Report* [*NASA Spec. Publ. 330* (no. 8), 1 (1973)].
10. R. R. Hodge, J. H. Hoffman, F. S. Johnson, in *Proceedings of the Fourth Lunar Science Conference* [*Geochim. Cosmoschim. Acta* **3** (Suppl. 4), 2855 (1973)]; J. H. Hoffman, J. R. Hodge, F. S. Johnson, *ibid.*, p. 2865.
11. A. E. Potter and T. H. Morgan, *Science* **241**, 675 (1988).
12. D. L. Anderson, *Earth Planet. Sci. Lett.* **18**, 301 (1973).
13. S. R. Taylor, *Lunar Science: A Post-Apollo View* (Pergamon, New York, 1975), p. 284.
14. H. J. Smith, in (*3*), pp. 37–42.
15. S. von Hoerner, in *Structures for Technology for Large Radio and Radar Telescope Systems*, J. W. Marr and H. Liebowitz, Eds. (MIT Press, Cambridge, MA, 1969), pp. 311–334. A study for a large lunar radio telescope of homologous design has been presented by F. Akyul, W. Gerstle, and S. W. Johnson, [in *Engineering, Construction, and Operations in Space II*, Proceedings of Space 90, S. W. Johnson and J. P. Wetzel, Eds. (American Society of Civil Engineers, New York, 1990), p. 697].
16. M. M. Colavita, M. Shao, D. H. Staelin, *Appl. Opt.* **26**, 4113 (1987).
17. P.-Y. Bely, C. J. Burroughs, G. D. Illingworth, Eds., *The Next Generation Space Telescope* (STSci, Baltimore, 1990).
18. J. N. Bahcall and L. Spitzer, Jr., *Sci. Am.* **247**, 40 (July 1982).
18a. *Space Science in the Twenty-First Century: Imperatives for the Decades 1995–2015* (National Academy Press, Washington, DC, 1988).
19. A. R. Thompson, J. M. Moran, G. W. Swenson, Jr., Eds., *Interferometry and Synthesis in Radio Astronomy* (Wiley-Interscience, New York, 1986).
20. A. R. Thompson *et al.*, *Astrophys. J. Supp. Ser.* **44**, 151 (1980).
21. F. R. Schwab, *SPIE J.* **231**, 18 (1980).
22. R. A. Perley, J. W. Dreher, J. Cowan, *Astrophys. J.* **285**, L35 (1984).
23. R. Braun, S. F. Gull, R. A. Perley, *Nature* **327**, 395 (1987); A. C. S. Readhead and P. N. Wilkinsin, *Astrophys J.* **223**, 25 (1978).
24. B. F. Burke, *SPIE-Reflective Opt.* **751**, 50 (1988).
25. ———, *Nature* **223**, 389 (1969).
26. M. M. Colavita, thesis, Massachusetts Institute of Technology (1985).
27. T. J. Cornwell, *IEEE Trans. Antennas Propag.* **36**, 1165 (1988).
28. R. Angel, personal communication.
29. R. Stachnik and A. Labeyrie, *Sky Telescope* **67**, 205 (1984).
30. See the review by J. A. Zensus and T. J. Pearson, in *The Impact of VLBI on Astrophysics and Geophysics*, M. J. Reid and J. M. Moran, Eds. (International Astronomical Union Symposium 129, Kluwer, Dordrecht, 1988), pp. 7–16.
31. A. Boss, B. F. Burke, M. J. Mumma, H. J. Smith, *Towards Observations of Planetary Systems (TOPS)* (NASA Conference Proceeding, NASA, Washington, DC, in press).
32. B. F. Burke, *Nature* **322**, 340 (1986); J. R. P. Angel, A. Y. S. Cheng, N. J. Woolf, *ibid.*, p. 341; B. F. Burke, *ibid.* **324**, 518 (1986); R. Angel, A. Cheng, N. J. Woolf, *ibid.*, p. 518.
33. T. Owen, in *Strategies for Search for Life in the Universe*, M. Papagianis, Ed. (Reidel, Dordrecht, 1980), pp. 177–185.
34. M. J. Rees, in *Quasars*, G. Swarup and V. Kapahi, Eds. (International Astronomical Union Symposium 119, Reidel, Dordrecht, 1986), p. 1.
35. This research was supported in part by grants from NASA and NSF. Many colleagues have contributed ideas and critiques; in particular, I thank R. Angel, D. Black, R. Brown, R. Burns, A. Layberie, M. Mumma, R. Stachnik, M. Shao, and N. Woolf. The LOUISA Workshop (*6*) in particular served as an important stimulus. The text is a development of a talk given at a Jet Propulsion Laboratory symposium, "Space Science of the Future," at the time of the Voyager encounter with Neptune.

Research Articles

Spacing Differentiation in the Developing *Drosophila* Eye: A Fibrinogen-Related Lateral Inhibitor Encoded by *scabrous*

NICHOLAS E. BAKER, MAREK MLODZIK, GERALD M. RUBIN

In the development of multicellular organisms a diversity of cell types differentiate at specific positions. Spacing patterns, in which an array of two or more cell types forms from a uniform field of cells, are a common feature of development. Identical precursor cells may adopt different fates because of competition and inhibition between them. Such a pattern in the developing *Drosophila* eye is the evenly spaced array of R8 cells, around which other cell types are subsequently recruited. Genetic studies suggest that the *scabrous* mutation disrupts a signal produced by R8 cells that inhibits other cells from also becoming R8 cells. The *scabrous* locus was cloned, and it appears to encode a secreted protein partly related to the β and γ chains of fibrinogen. It is proposed that the *sca* locus encodes a lateral inhibitor of R8 differentiation. The roles of the *Drosophila* EGF-receptor homologue (DER) and *Notch* genes in this process were also investigated.

THE STRUCTURE AND FUNCTION OF MOST MULTICELLULAR organisms depends on the differentiation of the cell types in an appropriate spatial organization. In some cases, it may be that the organization is determined by an invariant pattern of cell divisions. In others the organization is determined by induction, a process by which uncommitted cells are instructed to adopt a particular fate by another cell type. In a spacing pattern, some cells from a field of equivalent precursor cells become different from their neighbors, forming a spaced array of determined cells (Fig. 1). Which cells are chosen depends not on lineage or induction, but on the distance separating the differentiating cells, an indication that spacing patterns may result from competition; for example, the first of a set of equipotent cells to attain some developmental threshold inhibits the progression of others, giving rise to a mixture of determined and undetermined cells (Fig. 1).

Examples of spacing patterns are found in the development of many organisms. In insects, neuroblasts arise from the neurogenic ectoderm separated by undifferentiated cells that later make epidermis (*1*). Other examples include the pattern of sensory bristles in insects (*2*), determination of the anchor cell and the ventral uterine precursor cells in the nematode *Caenorhabditis elegans* (*3*), the differentiation of heterocysts in the nitrogen-fixing alga *Anabaena* (*4*), and the pattern of stomata on the leaves of higher plants (*5*). In vertebrates, the patterns of hair and tooth development may be examples of spacing patterns. Inhibitory mechanisms have been proposed to account for regeneration in Tubularia (*6*), vertebrate organ growth (*7*), and cell differentiation in the vertebrate retina (*8*). The wide phylogenetic distribution of spacing patterns suggests that competition and inhibition may be a common mechanism whereby

The authors are with the Howard Hughes Medical Institute and the Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

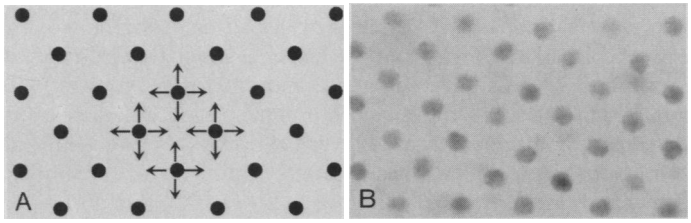

Fig. 1. Spacing patterns during development. The proposed role of lateral inhibition in spacing development is illustrated in (**A**). If differentiating cells inhibit their neighbors from following the same developmental pathway, a mixture of differentiated and undifferentiated cells can develop from an initially uniform population. The spacing between differentiating cells is determined by their field of inhibitory influence (represented by arrows). Experimental evidence for inhibition comes from ablation experiments. For example, when insect neuroblasts or nematode vulval precursors are ablated, a replacement is generated from the neighboring cells. Before ablation, differentiation of the replacement must have been inhibited by the original neuroblast or vulval cell (*51*). (**B**) R8 cells differentiating in the developing *Drosophila* retina (*52*). [Magnification ~300×.] About one-twentieth of the cells in the epithelium become R8 cells in a regular pattern. The formation of the regular array of R8 cells is thought to depend on a spacing pattern mechanism.

equivalent cells can produce a pattern of different cell types. However, the putative inhibitory factors that are essential to the competition-inhibition mechanism have not been identified.

We now describe mutations that affect a spacing pattern in the development of the *Drosophila* eye (Fig. 1). Our main conclusion is that the *scabrous* gene product, a putative secreted protein, is one of the factors used to prevent neighboring cells from entering the same developmental pathway. Mutations in the *Notch* and *DER* genes also affect this process.

Development of the eye imaginal disc in wild-type and *scabrous* mutant *Drosophila*. *Drosophila* eye development has been well reviewed (*9*). Each eye comprises about 800 identical ommatidia, or unit eyes, each containing 20 distinct adult cells (Fig. 2a). The eye forms from a single-cell thick epithelium, the eye imaginal disc. The R8 photoreceptor cell, which occupies a central position in each ommatidium, was identified by Tomlinson and Ready (*9*) as the first to begin differentiation; they inferred that specification of the R8 photoreceptor was the first step in ommatidium formation, and called this cell the "foundation photoreceptor." In each row, R8 cells are about eight cells apart, and the R8 cell array is thought to form as a spacing pattern (Fig. 1). Surrounding undifferentiated cells are then recruited over a 2-day period until each ommatidium achieves its full complement. Differentiation of the whole eye disc does not occur simultaneously. Instead, a wave of differentiation moves across the eye disc from posterior to anterior. The first events as the wave passes are the cessation of general cell division and a transient change in cell shape, so that a "morphogenetic furrow" separates undifferentiated from differentiating regions. This means that each eye disc contains ommatidia at many stages of development from the morphogenetic furrow to the posterior disc margin, arranged in the order of their formation (*10*).

Cell lineage plays no role in the determination of the eye disc cells, which are thought to be equivalent until cell division stops. The eye disc is not in contact with any other structure capable of inducing the R8 cell pattern, which can continue to develop if the disc is transplanted elsewhere in the larva (*11*). Therefore, the R8 cell array may arise de novo through a process of cell interactions within the epithelium (*12*). As with other spacing patterns, the spacing of the R8 cells may depend on mutual inhibition so that the distance of separation reflects the inhibitory sphere of influence of each R8 cell (*9*) (Fig. 1).

Mutations that alter the spacing of the R8 cells would be expected to alter the ommatidia in the adult. Many mutations have been described that affect adult eye morphology in *Drosophila* (*13*). Such mutations have been examined to identify those that changed the spacing of ommatidia in the eye disc. Three genes were found: *scabrous* (*14*); the *Drosophila* EGF-receptor homologue (*15*); and *Notch* (*16, 17*). Mutations at these loci are not specific to the eye and can also have effects on eye development after the R8 cell array is formed.

The sca^1 mutation is recessive (*18*) and, when homozygous, produces a roughening of the external surface of the eye. The number and spacing of adult bristles are also affected. Of the many other mutant alleles isolated, we use sca^{BP2} as representative because the sca^{BP2} allele is unable to synthesize any gene product (*19*). We refer to sca^{BP2} as sca^-. In sca^-, the eye is the same overall size as that in wild type and has a similar number of facets (the externally visible surfaces of the ommatidia), but the facets have irregular size and shape (Fig. 2, A and B). Sections of fixed material showed that ommatidia contained variable numbers of photoreceptor cells (Fig. 2, C and D). In a sample of 210 ommatidia, 40 percent were abnormal. The number of rhabdomeres in sections of abnormal ommatidia varied from 4 to 20 with a mean of 10 (*20*); in each section in wild type, 7 are seen.

In studies of the imaginal disc with antibodies that label specific ommatidial cells, the position of the ommatidia was abnormal as soon as they could be detected. Ommatidia were often closer together than in wild type, separated by fewer undifferentiated cells than normal (Fig. 2, E to H). Reduced separation of ommatidia was not due to death of the intervening cells because the extent of cell death was no greater in sca^- discs than in wild type (*21*). The separation between ommatidia could also be larger than normal. At

Table 1. Analysis of 591 *sca*-mosaic ommatidia, 327 for sca^{BP2} and 264 for sca^{OB7}, and 229 control mosaic ommatidia in which all cells were sca^+ (*50*). Ommatidia are not constructed by a pattern of cell divisions; thus clone boundaries take variable courses through individual ommatidia and any combination of mutant and wild-type cells is possible. (Upper) On average, it is expected that in a population of mosaic ommatidia any given cell type will be of each genotype in 50 percent of the cases. The sca^- R8 cells are significantly underrepresented in normally constructed mosaic ommatidia. (Lower) Cells in abnormally constructed ommatidia cannot be identified with certainty and were not analyzed directly. However, the number of abnormally constructed mosaic ommatidia was similar to the shortfall of normal ommatidia with sca^+ R8 cells (number of normally constructed mosaic ommatidia with sca^+ R8 cells minus the number with sca^- R8 cells, calculated from the data in the upper portion), suggesting that this was their origin.

Genetic composition of normally constructed ommatidia mosaic for *sca*

R cell	Mutant (percent)			
	sca^{BP2}	sca^{OB7}	Combined sca^-	Control
R1	45	42	44	48
R2	38	51	44	55
R3	40	54	46	52
R4	43	52	47	41
R5	42	45	43	47
R6	42	41	42	41
R7	51	38	45	44
R8	20	23	21	50

Fate of ommatidia containing sca^- R8 cells

Gene	Abnormally constructed mosaic ommatidia	Missing class of normally constructed ommatidia with sca^- R8 cells
sca^{BP2}	116	127
sca^{OB7}	100	88
Both alleles	216	215

later stages, when more cells had been recruited, improperly spaced clusters were often fused. The fused ommatidia contained more than the usual complement of photoreceptor cell precursors.

The unusual spacing of the differentiating ommatidia might be due to their formation at inappropriate locations. To investigate this, we used two markers for pattern formation prior to the expression of cell-type specific antigens. The cobalt sulfide method stains the apical borders of cells, and apical cell contacts are the first sign of ommatidia formation in the developing eye disc (*9*). In sca^- the clustering of cells was abnormal even in the morphogenetic furrow, before expression of neural antigens begins (Fig. 2J). Expression of the *sca* gene itself provides another marker of early pattern formation. In wild-type eye discs, *sca* antigen is first detected in clusters of cells preceeding the morphogenetic furrow, and remains in R8 cells for about 8 hours, disappearing as overt differentiation begins after the furrow (*19*) (Fig. 2K). We identified a mutant, sca^{UM2} (*14*), phenotypically similar to sca^- but that produces normal levels of *sca* antigen. In sca^{UM2}, the clusters of positive cells ahead of the furrow are abnormally spaced and often fused, and give rise to an abnormal pattern of *sca*-expressing R8 cell precursors in the morphogenetic furrow (Fig. 2L). These two experiments show that the spacing of ommatidium formation is abnormal in *sca* mutants, beginning ahead of the morphogenetic furrow about 8 hours before the expression of neural genes.

The role of the sca^+ gene in individual ommatidial cells was investigated by means of mosaic analysis. The cells in which the *sca* gene functions can be deduced from mosaic ommatidia that contain some mutant and some wild-type cells. Ommatidia in which some cells were sca^-/sca^- and the others were $sca^-/+$ were obtained at the borders of clones of homozygous sca^- cells. Some such ommatidia had normal morphology in the adult, others did not. Only cells in which sca^+ function is not required are expected to be sca^-/sca^- in phenotypically normal ommatidia. All the photoreceptors could be sca^- without affecting the structure of the ommatidium (Table 1). However, fewer normal ommatidia had mutant R8 cells (about 40 percent of controls).

The mosaic experiment showed that the sca^+ gene has a function in the R8 cell, but no role could be detected for the sca^+ gene in any other cells. This suggested that spacing depended on sca^+ function in the R8 cells, because R8 cell genotype was quantitatively sufficient to account for all the mutant ommatidia seen (Table 1). Consistent with this result, *sca* gene expression becomes limited to the R8 cell precursors (*19*). Function of sca^+ in disc cells whose fate in the adult is unknown (and which therefore could not be scored directly in the mosaic analysis), like the "mystery cells" (*9*) seems

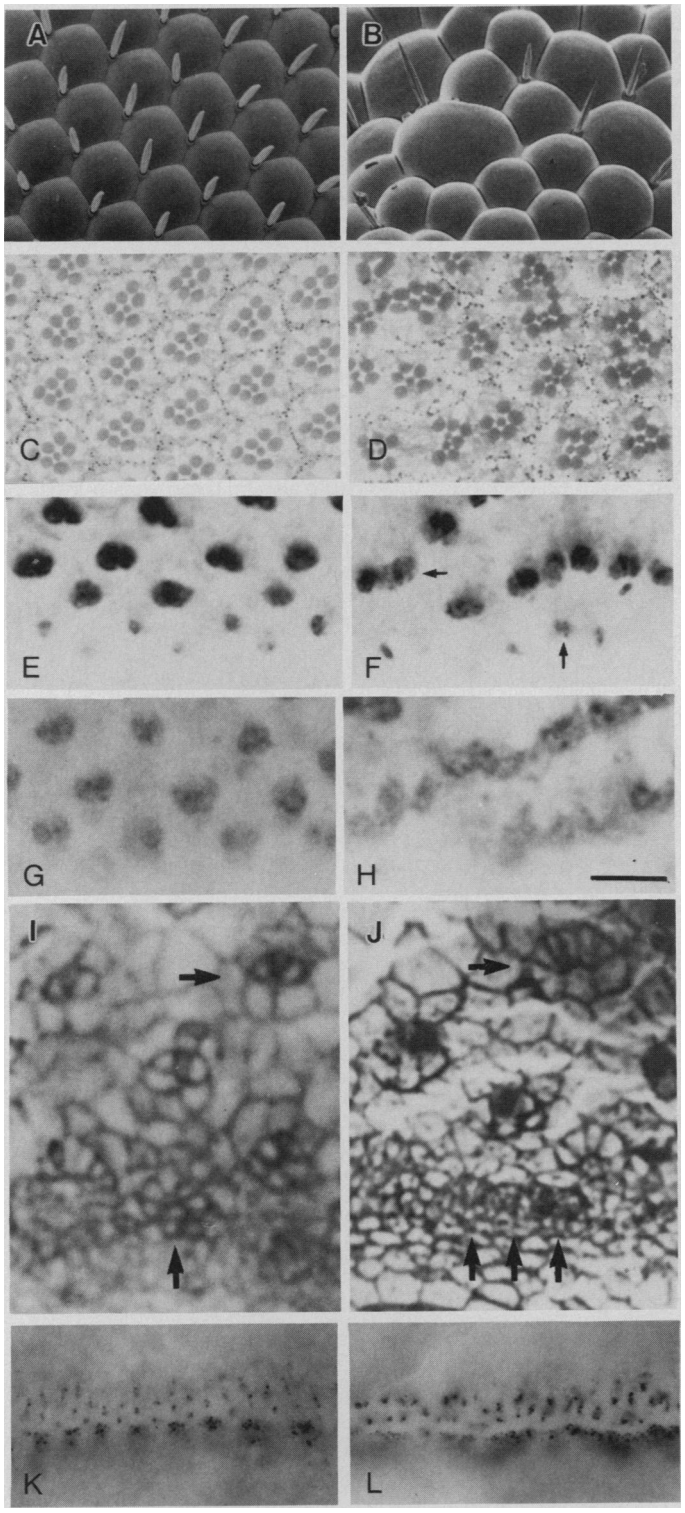

Fig. 2. The *scabrous* mutant phenotype. (**A**) Scanning electron micrograph of a portion of a wild-type *Drosophila* eye; about 20 of the ~800 facets are shown. (**B**) In a *sca* mutant eye (sca^{WB1}/sca^{WB1}) the facet size and hair positions are irregular (*53*). (**C**) In wild-type sections a regular arrangement is seen; each ommatidium is surrounded by pigment cells and contains seven photoreceptor cells visible at this plane of section (*53*). (**D**) In sca^-, pigment cells surround larger numbers of irregularly arranged photoreceptor cells. (**E**) Part of a wild-type disc stained with monoclonal antibody MAb22C10, which labels photoreceptor cells (*53*). The morphogenetic furrow is at the bottom, and older ommatidia are toward the top. Nearest the furrow, only R8 cells are positive; later three cells express the antigen. In the array, each row is exactly out of phase with the next. Eventually all the unlabeled cells are recruited to form part of the ommatidium such as the pigment cells that surround the neurons, or die (*9*). (**F**) sca^- eye disc. Note example of an ommatidium, apparently containing two R8-cells (vertical arrow), that is predicted to develop into an oversize cluster (horizontal arrow). (**G**) Wild-type eye disc stained with an antibody to the sevenless gene product. At this stage Sevenless is predominantly in R3 and R4 (*54*). (**H**) sca^- eye disc. (F and H) Examples of reduced spacing in one direction and increased spacing in another. Spacing can be reduced laterally, vertically, or diagonally (H), suggesting that interactions both within and between rows may be affected in sca^-. (**I**) Apical cell boundaries revealed by CoS staining in a wild-type eye disc (*53*). Ommatidia are first detected as clusters of cells in the morphogenetic furrow (vertical arrow). In more mature ommatidia (horizontal arrow) cells can be identified from their stereotyped positions. (**J**) In sca^- the ommatidia are irregularly spaced in the furrow (vertical arrows). More mature ommatidia often have too many cells (horizontal arrow). (**K**) In wild-type, *sca* is expressed briefly in clusters of cells ahead of the furrow and continues in R8 cells for about 8 hours. Antigen is perinuclear and in bodies which may be vesicular throughout the cell but especially apically (shown here) (*55*). (**L**) Expression of *sca* in the sca^{UM2} mutant. The scale bar in (H) represents 30 μm for (C to D), 20 μm for (E to H), 10 μm for (I to J), and 50 μm for (K and L).

Fig. 3. Deduced amino acid sequence of the *sca* protein. The pcs6 cDNA sequence has been deposited with GenBank (accession number M37703). Translation of the *sca* protein is expected to begin at nucleotide 299, and end with a TAA codon at nucleotide 2620. Intron positions, indicated by triangles above the protein sequence, were determined by comparison between the cDNA sequence and corresponding genomic regions (*56*). The algorithm of von Heijne (*57*) gives a score of +7.9 for signal peptide cleavage after Ala^{22} (scores above +3.5 are rare except for secreted proteins), and this putative signal sequence is underlined. Cleavage after Gly^{26} would also be possible (von Heijne score = +6.3). The putative Sca signal sequence lacks the basic NH_2-terminus (*n*-region) found in many signal peptides, but not in all (*58*). There are possible NH_2-linked glycosylation sites at residues 347, 562, 593, 635, 719, 761, and RR and KR dibasic residues: 294 to 296; 316 to 317; 468 to 469; 614 to 615; 709 to 710. There is an *opa* repeat which is mostly translated as threonine (dotted). In sca^{OB7} a G to A transition at nucleotide 816 replaces Trp^{173} with a TAG stop codon (asterisk). In sca^{WB1} a G to A transition at nucleotide 2291 replaces Trp^{698} with a TAG stop codon (asterisk) (*56*). Residues 504 to 695 of Sca are shown aligned with the corresponding regions from human fibrinogen β and γ chains, and a cDNA isolated from human adrenal medulla (*23, 24*). Within a 176–amino acid segment Sca has between 64 and 69 identities with these sequences (shaded), and therefore is almost as closely related to the fibrinogens as fibrinogens are to each other. Six residues are absolutely conserved between these and other fibrinogen-like sequences (circles), including four cysteines. The adrenal sequence is most similar (39 percent identity over the 176–amino acid residues (*24*). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

unlikely because more phenotypically abnormal mosaic ommatidia would be expected if another cell were also involved. In both sca^- mosaics and mutant homozygotes some ommatidia developed normally, suggesting that the function of *sca* is partially redundant and can sometimes be executed by other gene products.

The *scabrous* locus encodes a secreted protein related to fibrinogens. The *sca* gene has been isolated by transposon-tagging and the *sca* transcription unit has been identified (*19*). The cDNA clone pcs6, isolated from an eye-disc cDNA library, contains a 2791 base pair (bp) insert comprising nearly all the *sca* transcript (*22*). Conceptual translation reveals a single long open reading frame that extends from nucleotide 223 to 2619. The second ATG, at nucleotide 299, is likely to encode the amino-terminal methionine, because of its agreement with the known eukaryotic consensus for translation initiation and codon usage, making a predicted protein of 774 amino acids (Fig. 3). Hydropathy analysis defines only one hydrophobic region, extending from amino acids 1 to 35, with the most probable site for cleavage of a signal peptide being directly after Ala^{22}. The putative protein contains potential glycosylation sites and pairs of dibasic residues that are sometimes substrates for proteolytic enzymes (Fig. 3). These features suggest that the *sca* gene encodes a secreted protein.

Genomic DNA corresponding to the *sca* coding region was also sequenced from the sca^{OB7} and sca^{WB1} mutants. These mutations were induced by ethane methyl sulfonic acid (EMS) and are phenotypically similar to sca^- (*14*). In both alleles a G to A transition introduces a premature termination codon into the open reading frame (Fig. 3).

Comparison of the putative *sca* protein to the sequences of other known proteins revealed similarity between the carboxyl-terminal portion of the protein and fibrinogens, a class of secreted proteins in vertebrates (Fig. 3). The identities include four cysteines that are conserved in all fibrinogens and form intrachain cysteine bonds in these molecules. Several other proteins also contain a similar domain (*23*); a gene expressed in human adrenal medulla (*24*) has the greatest similarity to *sca*. We have been unable to detect significant sequence similarity to these or other known proteins outside the COOH-terminal region.

The similarity of parts of Sca and fibrinogens does not suggest an identical structure for these proteins. Vertebrate fibrinogens are secreted proteins that polymerize to form fibrin, a major constituent of blood clots (*25*). Fibrinogens form an αβγ-heterotrimer crosslinked by Cys-Cys and other covalent bonds, but none of the crosslinked regions are conserved in Sca, nor are regions important for other aspects of fibrinogen function, such as platelet binding. Fibrin polymerization is activated by cleavage of terminal fibrinopeptides. Neither fibrinopeptide-like sequences nor the polymerization sites themselves are present in Sca. Thus the putative *sca* expression product is unlikely to have a fibrinogen-like structure as a whole, and there is no reason to expect that it polymerizes. No biochemical functions have been localized to the domain conserved between fibrinogen and Sca, but Xu and Doolittle (*23*) have suggested that this region is involved in protein-protein interactions; therefore Sca may bind to some other protein or proteins.

Properties of the *sca* locus suggest a role in lateral inhibition. In the developing eye, formation of an array of ommatidia appears to depend on an inhibitory signal that prevents indiscriminate ommatidium formation, probably by specifying a regular array of cells as R8 precursors. We traced the origin of fused ommatidia found in sca^- mutant adults to defects in a prepattern that precedes the morphogenetic furrow. From this altered prepattern an abnormal array of R8 precursors appears. Our interpretation is that, when ommatidia form too close to each other, insufficient cells remain between them for further cell types to be recruited; later cell types can only be recruited around both primordia, resulting in the fused ommatidia seen in the adult.

Our results are most consistent with a role for the sca^+ gene in sending signals that regulate spacing in the eye. Fused ommatidia in sca^- indicated that a normal sca^+ function is to suppress formation of closely spaced R8 cells, and mosaic analysis attributed the mutant defect to loss of gene function in the R8 cells themselves. We assume that, when a sca^- R8 cell is formed in a mosaic region, inferior inhibition often allows ectopic ommatidia to form nearby. Thus ommatidia with sca^- R8 cells are more likely to fuse with misspaced neighbors, and be scored as phenotypically abnormal (*26*). By contrast, the alternative model that sca^+ function is required for receiving or interpreting an inhibitory signal is not consistent with the data in Table 1(*27*). Although the role of *sca* could be indirect,

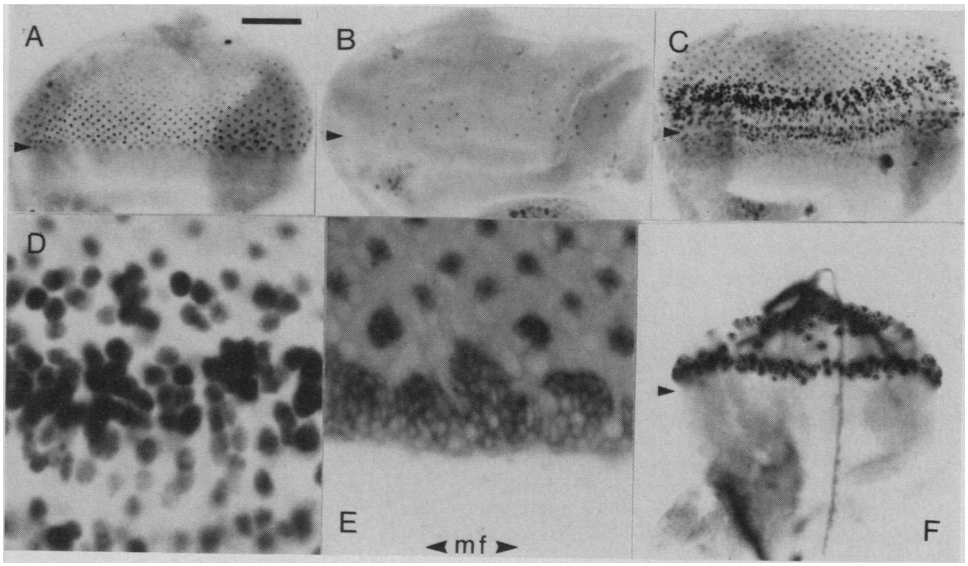

Fig. 4. Effects of *Elp* and N^{ts1} mutations on R8 spacing (**A**) Expression of β-galactosidase under the control of a *sca* enhancer (*52*) labels the R8 cells of wild-type discs. Shown here is a N^{ts1}/+ control disc after a temperature shift (6 hours at 32°C and then 12 hours at 18°C) (*52, 53*). In A, B, C, and F the scale bar represents 100 μm, and an arrow marks the position of the morphogenetic furrow. (**B**) In Elp^1/Elp^1 many fewer ommatidia develop (*15*). Their R8 cells express *sca*. (**C**) Expression of β-galactosidase under *sca* control in an N^{ts1}/Y fly after a temperature shift (6 hours at 32°C and then 12 hours at 18°C). A band of cells that were close to the morphogenetic furrow during the high-temperature treatment all contain β-galactosidase. In N^{ts1}/+ controls, β-galactosidase expression is the same as in wild type, as in (A). (**D**) Higher magnification of the disc from (C) showing β-galactosidase–containing cells. (**E**) Similar disc stained with MAbBP104, which labels all photoreceptors (*59*). The band of cells containing β-galactosidase is also differentiating as neurons. MAbBP104 is not expressed close to the morphogenetic furrow (mf), where neuronal differentiation has not begun. Similar results were obtained with MAb22C10, which detects another neuronal antigen (*9*). However, most of these cells did not express *sev* antigen, which is expressed in many cells but never in R8 (or R2 and R5) (*54*). (**F**) MAbBP104 staining in a N^{ts1}/Y; Elp^1/Elp^{B1} disc after a temperature shift (6 hours at 32°C and then 12 hours at 18°C). A band of cells close to the morphogenetic furrow differentiate as photoreceptors, although differentiation in the rest of the disc is reduced by the *Elp* mutation.

the locus has the capacity to encode a secreted protein so the most straightforward model is that the *sca* protein itself is an extracellular regulator of ommatidium spacing.

Other mutations affecting the array of ommatidia. Other gene products must also be required to regulate the pattern of R8 precursors. At a minimum, a receptor and proteins involved in signal transduction are expected. In addition, the fact that an aberrant pattern of ommatidia develops in sca^- rather than all cells developing identically suggests that *sca* might be only one of two or more partially redundant signals. Furthermore both activation and inhibition of R8 cell formation might be regulated. Thus a comprehensive picture of the generation of the R8 spacing pattern is not yet possible. However, two genes encoding transmembrane proteins, the DER gene and *Notch* (*N*), have been proposed to regulate ommatidium formation (*15, 16, 28*). We have performed a number of experiments aimed at elucidating how *sca* and these two loci might function during ommatidium formation.

In dominant mutations of the DER (*Elp* alleles) very few ommatidia develop (in *Elp/Elp*) and instead most cells remain undifferentiated (*15*). We looked at *sca* gene expression in *Elp/Elp* flies (Fig. 4, A and B). The *sca* gene was not active in the regions where ommatidia were absent, and was only expressed in the R8 cells of the few ommatidia that did form in *Elp/Elp* individuals. This indicated that no array of R8 cells was formed in the regions where no ommatidia develop, and that the *Elp* mutations probably affect eye development at a step that precedes the cell interactions involving the *sca* gene.

Notch protein expression (*29*) and the phenotype resulting from several *N* mutations suggest *N* functions in eye development (*30*). In studies of a temperature-sensitive mutation (N^{ts1}) Cagan and Ready (*16*) found that *N* activity was required to prevent indiscriminate ommatidium formation. A different mutation, *split* (N^{spl}) that results from a single amino acid change in the extracellular domain (*31*) recessively affects ommatidium formation and interacts genetically with *sca*. All *sca* alleles are dominant suppressors of N^{spl} so that the adult eye phenotype of N^{spl}/Y males (or N^{spl}/N^{spl} females) is suppressed if they are simultaneously *sca*/+. Duplications for the wild-type *sca* gene enhance the N^{spl} phenotype (*32*).

In the case of N^{ts1}, specific antibodies were used as markers to characterize the new photoreceptors that develop after reduction of N^+ function. A *lacZ* gene under the control of a *sca* enhancer expresses β-galactosidase in R8 cells (*19*), and other antibodies were used that label other cells. Judging from the expression of β-galactosidase and the other antigens, most of the new photoreceptor cells in N^{ts1} behave like R8 cells (Fig. 4, C to E). This suggested that in the wild type these cells all have the potential to become R8 but that N^+ activity suppresses this activity in all but a few. The effect of the N^{ts1} mutation in the presence of *Elp* mutations was also examined. Reduction of *N* function in the morphogenetic furrow region also increased the number of photoreceptor cells in an *Elp/Elp* imaginal disc (Fig. 4F), arguing that the DER mutation suppresses R8 formation by a mechanism requiring N^+ activity.

Comparison of eye development in a N^{spl}/Y male with wild type (Fig. 5) shows that fewer ommatidia develop in the mutant, and that they are spaced further apart than in wild type (*16, 17*). In the adult the rows of ommatidia are disordered and the eye is smaller than normal (*16, 17*). Suppression of the N^{spl} phenotype by *sca* mutations resulted from the restoration of a more normal pattern of ommatidial spacing (Fig. 5, F and H). In contrast there is no reciprocal interaction; N^{spl}/+ did not change the phenotype of sca^- homozygotes, and N^{spl}/Y; sca^- was similar to sca^- alone (Fig. 5G) (*33*). We used mosaic analysis to determine which cells caused the N^{spl} phenotype. When N^{spl}/N^{spl} clones were induced in N^{spl}/+ eyes, only about 40 percent of the expected number of mosaic ommatidia with N^{spl}/N^{spl} R8 cells were found, and the genotype of the other cells was not significant (*34*). Thus the absence of ommatidia is due to the effect of the N^{spl} mutant in the R8 cell precursors. In the N^{spl} mutant *sca* expression was not limited to developing ommatidia, but was also seen in regions from which ommatidia were absent (Fig. 5, D and E). Thus, in this genotype *sca* expression does not become restricted to differentiating R8 cells. This result contrasted with the effect of *Elp* mutations, which decreased both the number of ommatidia and *sca* expression coordinately (*35*).

The foregoing results show that *N* mutations also affect the spacing of R8 cells, and hence the *N* and *sca* genes may act at the same step. However the N^{ts1} and N^{spl} mutations have contrasting

effects. Reduction in function due to N^{ts1} gives more R8-like cells, but in N^{spl}/Y the number of ommatidia is decreased (*36*). Also, N^{spl} acts in the R8, but N^+ function should be dispensable for R8 determination because R8-like cells developed when N^+ activity was reduced. Therefore, the N^{spl} mutation may encode a protein with altered or increased activity, not reduced, and it is the inappropriate action of this product in the R8 cells that causes abnormal development (*37*).

Implications for spacing patterns in *Drosophila* and other organisms. We have used *Drosophila* eye development to look for genes involved in spacing patterns. On the basis of the phenotype of *sca* mutations and the sequence of the *sca* gene, we propose that the *sca* product acts as a lateral inhibitor in eye development. We think that *sca* expression, which in the eye is refined from groups of cells to R8 cells only (*19*) (Fig. 2K), reflects restriction of R8-forming potential to a few cells and then one, and we have shown that this refinement itself requires sca^+ function. Initially puzzling is that, in sca^-, reduced spacing between some ommatidia is accompanied by increased spacing between others. One explanation involves considering the influence of ommatidia in adjacent rows. In wild type, each row forms precisely out of register with the last so that ommatidia in the nascent row form maximally distant from existing ommatidia (Fig. 2, E and G). Reduced spacing in one row might therefore be expected to result in a local increase in spacing in the next (Fig. 2, F and H). Alternatively, theoretical models suggest that lateral inhibition should be coupled to lateral activation (*38*). In such models mutants in either process should also affect the other, and so may have dual phenotypes.

Previously it was thought the neurogenic genes might encode lateral inhibitory signals in *Drosophila*, but recent studies argued that *Notch* mutations usually act autonomously, and none of the neurogenic loci is yet known to encode a secreted protein (*39*). Also the increasing number of different developmental decisions known to require N^+ function may argue against an instructive role of *N* activity (*16, 39, 40*). The characterization of *sca* provides a novel component to be considered in models of neural development. Our studies support the idea that *N* is necessary to mediate the *sca* signal. A simplistic model, consistent with the genetic data, is that N might bind extracellular Sca. This is based on the N^{spl} point mutant in the *N* extracellular domain which makes R8 cells abnormally sensitive to *sca* product and the nonreciprocal genetic interaction (sca^- affects the phenotype of N^{spl}, but not vice versa) which suggests that information is transmitted in one direction only, from Sca to N. In tissue culture, N can mediate cell adhesion (*41*), and intracellular effects suggest it might act as a receptor (*35, 42*). At present there is no biochemical data on N-Sca binding and it is not known how N might transduce information to the inside of the cell.

Both *sca* and *N* are also required for development of the bristle pattern in *Drosophila* (*43*). Thus similar molecular mechanisms may be used to generate both these spacing patterns. An important difference concerns the role of the Achaete-Scute complex (AS-C). It is thought that bristle spacing is determined in two steps. First, a series of separate proneural regions are defined by expression of AS-C genes and their antagonists. Then lateral inhibition (requiring *N*, *sca*, and other genes) acts within each domain to refine the pattern to one or a few bristle mother cells (*19, 44*). In contrast, in the eye, both the formation of clusters of *sca*-expressing cells and their refinement to single R8 cells depend on *sca* and *N* function, and AS-C genes have no function in this process (*45*). Thus formation of the prepattern of *sca*-expressing clusters is not homologous to formation of proneural regions in bristle development. How a neurogenic band is formed and moves across the eye disc, and how lateral inhibition alone generates such a reproducible pattern, remain to be determined (*46*).

Less is known about spacing patterns in other organisms and whether similar proteins are involved. Other genes related to the whole *sca* molecule have not yet been reported, but the β/γ-fibrinogen domain is found in other proteins (*23*). No functions have been identified for any of these proteins except for the neural cell adhesion molecule cytotactin, which may be involved in neuronal-glial interactions during development (*47*). However, genes similar to *N* have been isolated from *C. elegans*, *Xenopus*, humans, and rats (*48*). In *C. elegans* the *N* cognate *lin-12* is involved in interactions between two cells that compete to become the Vu

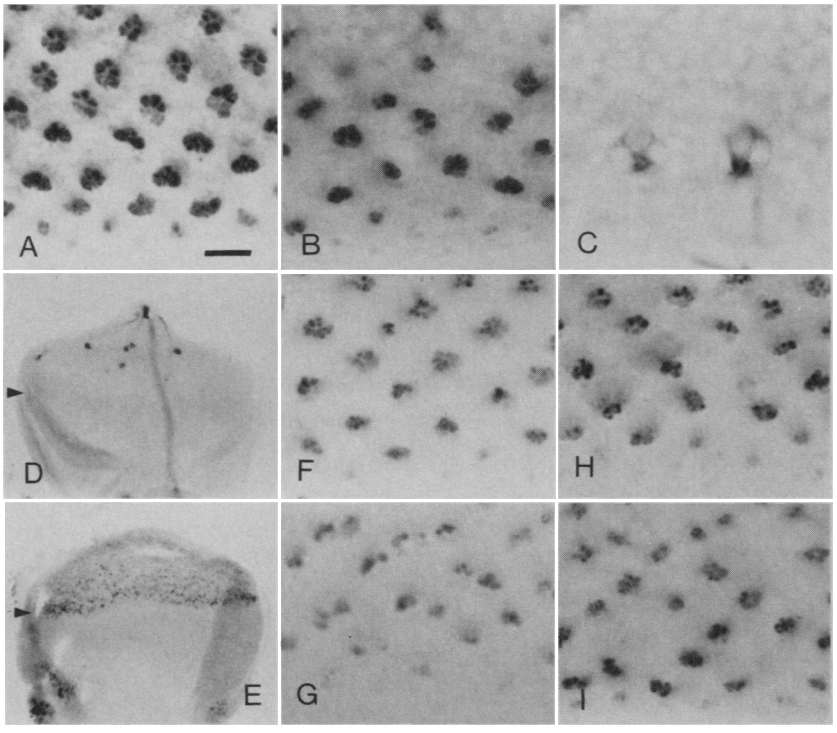

Fig. 5. Neuronal differentiation in the N^{spl} mutation. Photoreceptor differentiation in the imaginal discs of various genotypes as revealed by staining with MAb22C10, except (D) where MAbBP104 was used (*53,59*). (**A**) The normal pattern of photoreceptor cell differentiation (N^+/N^+). Scale bar represents 10 μm in A to C and F to I. N^{spl}/+ females are similar, but there are rare gaps in the pattern. (**B**) N^{spl}/Y male. (**C** and **D**) Very few ommatidia develop in N^{spl}/Y;$E(spl)^D$/+ (*28, 37*). (**E**) Expression of β-galactosidase under *sca* control in *spl*/Y; $E(spl)^D$/+. Unlike *Elp* (Fig 4B) abundant *sca* expression began in the morphogenetic furrow (arrowhead) although very few ommatidia develop (D). *sca* is also abnormally expressed in N^{spl}/Y in the absence of the $E(spl)^D$ mutation (data not shown). (**F**) An almost normal pattern of differentiation was seen in N^{spl}/Y; sca^{BP2}/+ males (compare to panels A and B). (**G**) N^{spl}/Y; $sca^{BP2}/Df(2R)sca^{KX2}$ disc. The pattern resembles sca^- (Fig. 2, E and F). (**H**) *sca* mutations suppress the N^{spl}/Y; $E(spl)^D$/+ phenotype. Shown here is N^{spl}/Y; Df(2R)sca^{KX2}/+;$E(spl)^D$/+. Compare to panel C. (**I**) N^{spl}/Y; DpN^+; $E(spl)^D$/+. An extra dose of the wild-type N^+ gene on Dp(1;2)51b6,N^+ w^+ suppressed the N^{spl}; $E(spl)^D$/+ phenotype (C). Similar suppression by the N^+ duplication was also seen in the following three genotypes: N^{spl}/Y, N^{spl}/N^{spl}, and N^{spl}/+; $E(spl)^D$/+.

precursor cell, and between six VPC cells. In each case the successful cell inhibits the others, strikingly reminiscent of R8 development in *Drosophila* (*49*). Thus, at least some spacing patterns in other organisms may be regulated by proteins similar to those used in *Drosophila*.

REFERENCES AND NOTES

1. V. Hartenstein and J. A. Campos-Ortega, *Roux's Arch. Dev. Biol.* **193**, 308 (1984).
2. V. B. Wigglesworth, *J. Exp. Biol.* **17**, 180 (1940).
3. J. Kimble and D. Hirsh, *Dev. Biol.* **70**, 396 (1979).
4. G. E. Fogg, *Ann. Bot.* (New Series) **13**, 241 (1949); M. Wilcox, G. J. Mitchison, R. J. Smith, *J. Cell Sci.* **12**, 707 (1973).
5. T. Sachs, in *The Clonal Basis of Development*, S. Subtelny and I. M. Sussex, Eds. (Academic Press, New York, 1978).
6. T. H. Morgan, *J. Exp. Zool.* **1**, 587 (1904).
7. P. J. Bryant and P. Simpson, *Q. Rev. Biol.* **59**, 387 (1984).
8. T. A. Reh, *J. Neurosci.* **7**, 3317 (1987).
9. A. Tomlinson and D. F. Ready, *Dev. Biol.* **120**, 366 (1987); A. Tomlinson, *Development* **104**, 183 (1988); D. F. Ready, *Trends Neurosci.* **12**, 102 (1989).
10. For simplicity we use "ommatidia" to describe all stages of ommatidial assembly, from the first preclusters that contain six cells to the mature 20-cell units.
11. B. Ephrussi and G. W. Beadle, *Am. Nat.* **70**, 218 (1936).
12. An array of ommatidia still forms after the existing pattern has been disrupted by mutations [G. G. Foster and D. T. Suzuki, *Proc. Natl. Acad. Sci. U.S.A.* **67**, 738 (1973); C. A. Poodry, L. Hall, D. T. Suzuki, *Dev. Biol.* **32**, 373 (1973); also see (*15*)] or irradiation [H. J. Becker, *Z. Induk. Abst. Vereb. Lehre* **88**, 333 (1957)]. Also transplantation experiments show a normal array can still form if the differentiated rows are removed [R. M. Lebowitz and D. F. Ready, *Dev. Biol.* **117**, 663 (1986)].
13. D. F. Lindsley and E. H. Grell, *Genetic Variations of* Drosophila melanogaster (Carnegie Inst. Wash. Publ. 627, 1968).
14. N. E. Baker, K. Moses, D. Nakahara, G. M. Rubin, in preparation.
15. N. E. Baker and G. M. Rubin, *Nature* (London) **340**, 150 (1989).
16. R. Cagan and D. F. Ready, *Genes Dev.* **3**, 1099 (1989).
17. P. J. Renfranz and S. Benzer, *Dev. Biol.* **136**, 411 (1989).
18. Discovery by P. T. Ives of sca^1, the first mutation at the *sca* locus, was reported in T. H. Morgan *et al.*, *Carnegie Inst. Wash. Year Book* **37**, 304 (1938).
19. M. Mlodzik, N. E. Baker, G. M. Rubin, *Genes Dev.* **4**, 1848 (1990). The sca^{BP2} chromosome contains a 2-kb deletion at the 5′ end of the gene. The first exon, including the transcription start site and beginning of the open reading frame, is deleted, and sca^{BP2} homozygotes make no detectable mRNA or protein.
20. The numbers of rhabdomeres per ommatidium did not fit a normal distribution. Instead frequencies clustered around 7 (wild type) and lesser maxima at 5, 9, and 14.
21. Cell death was assayed by basic fuchsin staining, essentially as described [E. Wieschaus and C. Nüsslein-Volhard, in *Drosophila: A Practical Approach*, D. Roberts, Ed. (IRL Press, Washington, DC, 1986), chap. 9].
22. The 5′ terminus of the pcs6 insert is 101 bp downstream from the transcription start site determined in primer-extension experiments, and includes a putative poly(A) addition sequence 27 nucleotides before its 3′ end (*19*).
23. X. Xu and R. F. Doolittle, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2097 (1990).
24. Y. Morel, J. Bristow, S. E. Gitelman, W. L. Miller, *ibid.* **86**, 6582 (1989).
25. R. F. Doolittle, *Annu. Rev. Biochem.* **53**, 195 (1984).
26. Another feature of abnormally constructed ommatidia was consistent with this view. Within normal ommatidia, position is a certain guide to cell identity, but R8 cells also have distinctive morphology. Their rhabdomeres lie centrally (R7 cells are the other central photoreceptors), and extend only to the basal part of the retina. In several abnormally constructed ommatidia we observed two cells with R8-like morphology, suggesting that the cluster formed around 2 R8 cells. One of these was always sca^-, as expected if the mutant phenotype results from loss of *sca* function in R8 cells.
27. In the cell autonomous model sca^- cells are more likely than sca^+ cells to become R8 cells and found an ommatidium. This would not reduce the ratio of sca^- R8 cells among ommatidia that developed normally.
28. J. A. Campos-Ortega and E. Knust, *Roux's Arch. Dev. Biol.* **198**, 286 (1990).
29. S. Kidd, M. K. Baylies, G. P. Gasic, M. W. Young, *Genes Dev.* **3**, 1113 (1989); K. M. Johansen, R. G. Fehon, S. Artavanis-Tsakonas, *J. Cell Biol.* **109**, 2427 (1990).
30. R. W. Pilkington, *Proc. Zool. Soc. London Ser. A* **111**, 199 (1942); W. J. Welshons, *Science* **150**, 1122 (1965); G. G. Foster and D. T. Suzuki, *Proc. Natl. Acad. Sci. U.S.A.* **67**, 738 (1970); D. L. Shellenbarger and J. D. Mohler, *Genetics* **81**, 143 (1975); U. Dietrich and J. A. Campos-Ortega, *J. Neurogenet.* **1**, 315 (1984).
31. Deficiencies or alleles that unconditionally eliminate *N* gene function are recessive lethal mutations because of effects on the embryonic nervous system. *N* is nevertheless symbolized by a capital letter because of a subtle dominant phenotype (wing morphology is affected) that is due to absence of one dose of the gene in a *N*/+ heterozygote. The *split* mutation results from the substitution of a isoleucine for a threonine residue in EGF-repeat number 14 [M. R. Kelly, S. Kidd, W. A. Deutsch, M. W. Young, *Cell* **51**, 539 (1987); D. A. Hartley, T. Xu, S. Artavanis-Tsakonas, *EMBO J.* **6**, 3407 (1987)] and is usually abbreviated *spl*. For the work described here, we use N^{spl} to emphasize that *split* is in fact a point mutation at the *N* locus.
32. M. Brand and J. A. Campos-Ortega, *Roux's Arch. Dev. Biol.* **198**, 275 (1990); L. Rabinow and J. A. Birchler, *Genetics* **125**, 41 (1990).
33. The homozygous sca^- eye phenotype also was not modified in N^-/+ females.
34. N^{spl} w^{1118}/+ females were irradiated (1000 rad, late first instar), and phenotypically w^- clones were sectioned and mounted for microscopic examination of ommatidium construction. The *w* genotype of individual cells was scored in 205 phenotypically normal N^{spl}-mosaic ommatidia. N^{spl}/N^{spl} R8 cells were found in only 41 (20 percent), but frequencies for other N^{spl}/N^{spl} photoreceptors was close to the expected 50 percent (R1, 48 percent; R2, 40 percent; R3, 43 percent; R4, 42 percent; R5, 41 percent; R6, 46 percent; R7, 41 percent. R2, R3, R4, and R5 are slightly less common; these cells are determined before the final mitosis and are therefore more likely to be related in lineage to R8). From these data we estimate that 105 more ommatidia with N^{spl}/N^{spl} R8 cells were expected (number of normally constructed mosaic ommatidia with N^{spl}/+ R8 cells minus the number with N^{spl}/N^{spl} R8 cells). Only 28 abnormally constructed mosaic ommatidia were found, implying that most ommatidia with N^{spl}/N^{spl} R8 cells are not formed or fail to develop.
35. In N^{spl}; $E(spl)^D$/+, *sca* expression, as reflected by *lac Z* expression driven by a *sca* enhancer (*52*), does not correlate with R8 differentiation. This suggests that N^{spl} affects R8 development after *sca* expression. However, the N^{ts} experiment argued that *N* acts upstream of *sca* expression. Therefore it appears that *N* functions both upstream and downstream of *sca*. Feedback mechanisms are likely to be important in the competition to make the R8 cell pattern, as our data on the dependence of the *sca* expression pattern on *sca* function shows (*19, 38*). *N* may be a component of the feedback mechanism.
36. Our results cannot distinguish whether N^{spl} affects the determination of the R8 precursors, or if N^{spl} R8 cells are determined, but are then unable to continue normal development. In N^{spl}/Y, cell death occurs close to the morphogenetic furrow (*16, 28*, and N. E. Baker, unpublished data). Thus cell death might be responsible for early loss of N^{spl} ommatidia that form in a normal pattern. Alternatively, if N^{spl} affects R8 determination, cell death might occur as a consequence of the altered pattern.
37. Mutations that alter or increase gene activity are expected to be dominant, but N^{spl} is often described as recessive. In fact the situation is more complicated. N^{spl} has dominant effects on bristles in the thorax [C. Stern and C. Tokunaga, *Proc. Natl. Acad. Sci. U.S.A.* **60**, 1252 (1968)] and abdomen (*13*), in eye development in N^{spl}/+ females (N. E. Baker and G. M. Rubin, unpublished data) and in the background of the mutations *Enhancer of split* [$E(spl)^D$; W. J. Welshons, *Drosoph. Inf. Serv. Bull.* **20**, 157 (1956)] or *Ellipse* [N. E. Baker and G. M. Rubin, unpublished data]. These are not loss-of-function phenotypes because they are not seen in Df*N*/+ flies. Since N^{spl} can be dominant or recessive depending on the tissue and the genetic background it is probable the N^{spl} protein interacts with other products, and that these interactions determine whether N^{spl} is dominant or recessive in a particular circumstance. Genetic and biochemical data suggest that *N* molecules can dimerize and that different *N* molecules can modulate each other's function (*41*) [T. Xu, I Rebay, R. J. Fleming, T. N. Scottdale, S. Artavanis-Tsakonas, *Genes Dev.* **4**, 464 (1990)]. Extra copies of the N^+ alleles always supress N^{spl} phenotypes in both adults and eye imaginal discs (Fig. 5). Thus it may be that N^{spl} has a gain or change in N^+ function that is often recessive because of interactions with the N^+ homologue. A similar interpretation has been proposed for the *Abruptex* alleles of *N* [J. Palka, M. Schubiger, H. Schwaniger, *Development* **109**, 167 (1990)].
38. A. M. Turing, *Philos. Trans. R. Soc. London Ser. B* **641**, 37 (1952); H. Meinhardt, *Mechanisms of Biological Pattern Formation* (Academic Press, New York, 1982).
39. R. J. Greenspan, *New Biol.* **2**, 595 (1990).
40. J. R. Jacobs, Y. Hiromi, N. H. Patel, C. S. Goodman, *Neuron* **2**, 1625 (1989).
41. R. G. Fehon *et al.*, *Cell* **61**, 523 (1990).
42. C. V. Cabrera, *Development* **110**, 733 (1990).
43. In addition to eye development, *sca* mutations affect the number and position of sensory bristles in the adult, which are also organs affected by *N* mutations. The N^{spl} bristle phenotype also depends on *sca* dosage, so *sca* might also encode a ligand for *N* during bristle development. Consistent with this hypothesis, the bristle phenotype of *sca* mutations is enhanced by heterozygosity for an *N* deletion (N. E. Baker *et al.*, unpublished results). See (*19*) for discussion of the possible role of *sca* during embryonic development.
44. P. Simpson, *Development* **109**, 509 (1990).
45. F. Jimenez and J. A. Campos-Ortega, *Roux's Arch. Dev. Biol.* **4**, 179 (1987).
46. Presumably some of the factors involved in activating neurogenesis may be modified in *Elp* mutants, as *sca* expression is reduced in *Elp*/*Elp*.
47. K. L. Crossin, S. Hoffman, S. Tan, G. M. Edelman, *Dev. Biol.* **136**, 381 (1989).
48. J. Yochem and I. S. Greenwald, *Cell* **58**, 553 (1989); C. Coffman, W. Harris, C. Kintner, *Science* **249**, 1438 (1990); L. Ellison and J. Sklar, personal communication; G. Weinmaster, personal communication.
49. I. S. Greenwald, P. W. Sternberg, H. R. Horvitz, *Cell* **34**, 435 (1983); P. W. Sternberg, *Nature* **335**, 551 (1988); P. W. Sternberg and H. R. Horvitz, *Cell* **58**, 679 (1989); G. Seydoux and I. S. Greenwald, *ibid.* **57**, 1237 (1989).
50. Mosaics were generated by x-ray–stimulated mitotic recombination. A copy of the *white* gene, transposed onto the second chromosome close to the *sca* locus (at 49D), was used as a dominant genetic marker. w^{1118}; $sca^{BP2}/P[w^+]49D$ or w^{1118}; $sca^{OB7}/P[w^+]49D$ larvae were irradiated (1000 rad, late first instar), as were w^{1118}; $sca^+/P[w^+]49D$ controls. Eyes containing w^- (and therefore homozygous for the sca^- experimental or sca^+ control chromosomes) clones were processed (*34*).
51. C. Q. Doe and C. S. Goodman, *Dev. Biol.* **111**, 206 (1985); J. E. Sulston and J. G. White, *ibid.* **78**, 577 (1980); J. Kimble, *ibid.* **87**, 286 (1981).
52. Strains in which β-galactosidase is expressed under the control of *sca* enhancers have been described (*19*). Line A2-6 was used in these experiments. The nuclear β-galactosidase protein was detected with a monoclonal antibody (Promega), and horseradish peroxidase–coupled goat antiserum to mouse IgG (Biorad), essentially as described (*9*).
53. Methods were as described in (*9*) and (*15*).

54. A. Tomlinson, D. D. L. Bowtell., E. Hafen, G. M. Rubin, *Cell* **51**, 143 (1987).
55. Polyclonal antisera specific for *sca* products were obtained after immunizing BALB/c mice with a TrpE-Sca fusion protein [T. J. Koener, J. E. Hill, A. M. Myers, A. Tzagoloff, *Methods Enzymol.*, in press] lacking only the NH_2-terminal 19 amino acids of the putative *sca* signal peptide.
56. pcs6 DNA was sequenced on both strands with the shotgun cloning method and chain-terminating inhibitors [A. T. Bankier and B. G. Barrell, in *Techniques in the Life Sciences, B5, Nucleic Acid Biochemistry* (B508), 1-34, R. A. Flavell, Ed. (Elsevier, New York, 1983)], and the modified T7 polymerase "Sequenase" [S. Tabor and C. C. Richardson, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4767 (1987)]. Oligonucleotides derived from the cDNA sequence and its complement were used as primers to sequence corresponding regions of a double-stranded genomic DNA template. Introns were identified from the discontinuity of the cDNA and genomic sequences. The intron-exon structure obtained was consistent with the cDNA and genomic restriction maps, with the distribution of genomic restriction fragments complementary to cDNA in hybridization experiments (see *19* for the genomic organization of the *sca* exons), and with the consensus sequences for splice donor and acceptor sites [S. M. Mount, *Nucleic Acids Res.* **10**, 459 (1982)]. Mutant DNA's were obtained by means of the polymerase chain reaction. Genomic DNA from *mutant/deficiency* or from wild-type strains was amplified with oligonucleotide primers corresponding to sequences flanking the four exons. The double-stranded reaction products were either sequenced directly with oligonucleotide primers, or cloned into m13 and sequenced from single-stranded recombinant bacteriophage DNA. Sequence differences were confirmed by sequencing products from more than one independent polymerase chain reaction.
57. G. Von Heijne, *Nucleic Acids Res.* **14**, 4683 (1986).
58. At least one other *Drosophila* protein, the product of the *wingless* gene, is known to be secreted despite lacking the n-region [F. R. Rijsewijk *et al.*, *Cell* **50**, 649 (1987); M. Van den Heuvel, R. Nusse, P. Johnston, P. A. Lawrence, *ibid.* **59**, 739 (1989)].
59. M. Hortsch, A. J. Bieber, N. H. Patel, C. S. Goodman, *Neuron* **4**, 697 (1990).
60. Supported in part by postdoctoral fellowships from the Damon Runyon–Walter Winchell Cancer Research Fund (N.E.B.) and EMBO (M.M.). We thank the laboratories of S. Artavanis-Tsakonas, J. Campos-Ortega, and L. Y. and H. N. Jan for *Drosophila* stocks and unpublished information, M. Simon for the negative printed in Fig. 2A, E. Bier, R. Cagan, R. Doolittle, C. Goodman, and P. Hoppe for interesting discussions, and R. Carthew, E. Ferguson, J. Fischer, M. Freeman, C. Goodman for comments on the manuscript. Sequences are available from GenBank under the accession number M37703.

10 July 1990; accepted 26 October 1990

How Big Is the Universe of Exons?

ROBERT L. DORIT, LLOYD SCHOENBACH, WALTER GILBERT

If genes have been assembled from exon subunits, the frequency with which exons are reused leads to an estimate of the size of the underlying exon universe. An exon database was constructed from available protein sequences, and homologous exons were identified on the basis of amino acid identity; statistically significant matches were determined by Monte Carlo methods. It is estimated that only 1000 to 7000 exons were needed to construct all proteins.

MOST GENES IN COMPLEX EUKARYOTES CONSIST OF short exons separated by long introns. In one view, genes are assembled, via intron-mediated recombination, from exon modules that code for functional domains, folding regions, or structural elements (*1*, *2*). Such models portray introns as a retained primitive feature. Alternatively, the phylogenetic distribution of introns has led to arguments that introns are a derived feature of eukaryotic genomes, the result of bursts of parasitic elements invading early (and continuous) eukaryotic coding regions (*3*, *4*).

The hypothesis of exon shuffling proposes that complex genetic information is built up by joining previously independent exons, thus giving rise to more complex proteins and to novel enzymatic functions. This view of the modular assembly of extant genes is supported by the common structural features of certain large gene superfamilies, such as the immunoglobulin-like superfamily (*5*), and by the examples of exon reuse observed in the mosaic structure of the LDL (low density lipoprotein) receptor and the EGF (epithelial growth factor) precursor (*6*). In other gene superfamilies, the older intron-exon gene structure is still apparent in certain representatives, while other members of the same family have lost introns (possibly through retroposition of a mature message) to produce genes with longer and more complicated exons, but with few or no remaining introns. An example of this pattern is the opsin superfamily, which includes genes with four introns as well as genes for beta-adrenergic receptors, which have no introns at all (*7*).

The ancient character of introns is also supported by data suggesting that introns antedate the divergence of plants and animals a billion years ago (*8*). Intron-exon structures may also predate the endosymbiotic incorporation of chloroplasts and mitochondria, which occurred about 2 billion years ago (*9*). Introns may, in fact, antedate the first branchings of life on Earth: the first protogenes may have already displayed intron-exon structure. The original exons may have been 15 to 20 amino acids long; processes of intron sliding and intron loss leading to more complex exons have produced the present day spectrum (*2*).

In this article, the frequency of exon shuffling events is surveyed in order to address the following question: How many different exons were required to generate the current protein diversity? We have identified homologous exons (those of common evolutionary origin) on the basis of amino acid sequence similarity. To the extent that every exon in an underlying universe of exons has an equal probability of being incorporated into a gene, we can then estimate the size of that underlying universe by determining how frequently homologous exons appear in nonhomologous genes.

We first constructed a database of all known exons. The available databases contain large numbers of homologous gene sequences; we eliminated such duplication in order to obtain a collection of exons derived solely from independent genes, unrelated by direct descent. We then made pairwise comparisons of all these independent exons to identify statistically significant sequence similarities, which, we argue, indicate exon homology.

Finally, using a simple sampling model, we took this number of exon repeats to estimate the size of the exon universe. If we survey n exons that have been drawn with replacement from an underlying set of size N, we expect the number of repeats to be given by the product of $n(n-1)/2$, the number of pairs of objects in the collection, and the probability that any pair will match, $1/N$. The number of single repeats is thus $n(n-1)/2N$. Accordingly the expectation for triple repeats is $n(n-1)(n-2)/6N^2$ and so forth (*10*).

The authors are with the Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138.