EMERGING TECHNIQUES

Sticky molecules in not-so-sticky cells

Michael Hortsch and Allan J. Bieber

The assignment of specific roles to cell-surface proteins by standard methods can be a major problem. In the technique described below, Schneider-2 (S2) cells, an established *Drosophila* cell line, have been used in cell transfection and aggregation experiments. As such, they have proved to be a useful tool for the functional characterization of putative cell-adhesion molecules.

THE FUNCTIONAL ANALYSIS of cellular proteins that have been identified by immunological or genetic approaches is a major problem in modern biology. A mutant phenotype or a specific antibody-staining pattern does not usually supply the information required to deduce the molecular function or mode of action of the protein affected by the mutation or recognized by the antibody. Unless the amino acid sequence has sufficient homology to other proteins with a known biological function, the analysis of a cloned cDNA sequence frequently does not provide any clues about the function of a protein. The expression of cDNAs encoding several different putative vertebrate cell-adhesion proteins in transfected tissue culture cells has been pioneered in the laboratories of Takeichi and Edelman to study the functional aspects of this type of molecule[1,2]. More recently, *Drosophila melanogaster* Schneider-2 (S2) cells have also been used as a powerful tool to study a number of *Drosophila* molecules for which such a biological function was suspected but had not yet been demonstrated.

S2 cells as a host for 'sticky' molecules

Within the past year a cell transfection system using an established *Drosophila* cell line and several different *Drosophila* expression vectors has been introduced. This system has provided clear evidence for a cell-adhesion and receptor function for a number of cell-surface glycoproteins that are expressed by different tissues at various times during *Drosophila* development. These molecules had been identified, isolated and cloned using either monoclonal antibodies or by molecular genetic approaches. Although their involvement in cell adhesion, recognition and signaling events had been postulated, there was no direct experimental evidence to support this hypothesis.

M. Hortsch is at the Department of Anatomy and Cell Biology, University of Michigan, Medical Science II Building, Ann Arbor, MI 48109-0616, USA. **A. Bieber** is at the Department of Biological Sciences, Purdue University, Lilly Hall of Life Sciences, West Lafayette, IN 47907, USA.

The centerpiece of the new cell-adhesion assay is the S2 cell line. This *Drosophila* cell line was established more than 20 years ago from primary cultures of 20–24 h embryos[3]. S2 cells are grown at 25°C in insect tissue culture medium (e.g. Schneider's, M3 or Grace's medium) supplemented with fetal calf serum and antibiotics without the need for a CO_2-controlled environment. Most cells grow in suspension as single, unattached cells with a rounded morphology and a doubling time of about 24 h. Only a few cells become loosely attached to the tissue culture support. The tissue origin of the S2 cell line is unclear, but in the original article describing the S2 cells, Schneider suggested that it might be derived from imaginal disc cells[3]. Their low endogenous adhesiveness and the ease with which they can be propagated as a continuous cell line make S2 cells ideal for cell transfection and cell aggregation experiments.

The general approach for testing a cDNA encoding the complete open reading frame of a putative cell-adhesion molecule is outlined in Fig. 1. After subcloning into a suitable expression vector and introduction into S2 cells by either calcium phosphate precipitation or some other transfection method, the cDNA can be expressed by activation of the promoter used, e.g. a heat shock treatment or the addition of Cu^{2+}-ions. Transfected cells can then be scored for adhesive properties by gentle agitation on a shaking platform and the aggregation process can be observed with a standard light microscope.

Several different expression vectors are available for high-level expression of proteins from their respective cDNAs. The pHT-4 vector described by Schneuwly *et al.*[4] and the pCaSpeR-hs vector (a derivative of the Carnegie 4 vector[5]) permit the regulation of cDNA expression by the hsp70 heat shock promoter. The pCaSpeR-hs vector, in contrast to the pHT-4 plasmid, has several different unique endonuclease restriction sites in its polylinker that can be used for the subcloning of DNA fragments. Both vectors are also suitable for transposase-mediated germ line transformation of *Drosophila* embryos. Alternatively, cDNAs can be put under the control of the *Drosophila* metallothionein promoter by subcloning into the pRmHa-3 vector[6]. mRNA transcription is induced by culturing the transfected cell line for 1–3 days in medium containing a low concentration of Cu^{2+}-ions (0.7 mM). Cu^{2+}-ions will not induce a heat shock response in S2 cells and might therefore be preferable if the expressed polypeptide is unstable under heat shock conditions.

A number of different selectable markers have been used for the isolation of transfected *Drosophila* cell lines. Vector DNA is usually introduced into S2 cells by cotransfection with the pPC4 plasmid. The pPC4 DNA contains an α-amanitin-resistant RNA polymerase II large subunit from *Drosophila* and transfected cells can be selected by culturing in 5 μg ml^{-1} α-amanitin[7]. Other groups used methotrexate or neomycine resistance markers in their

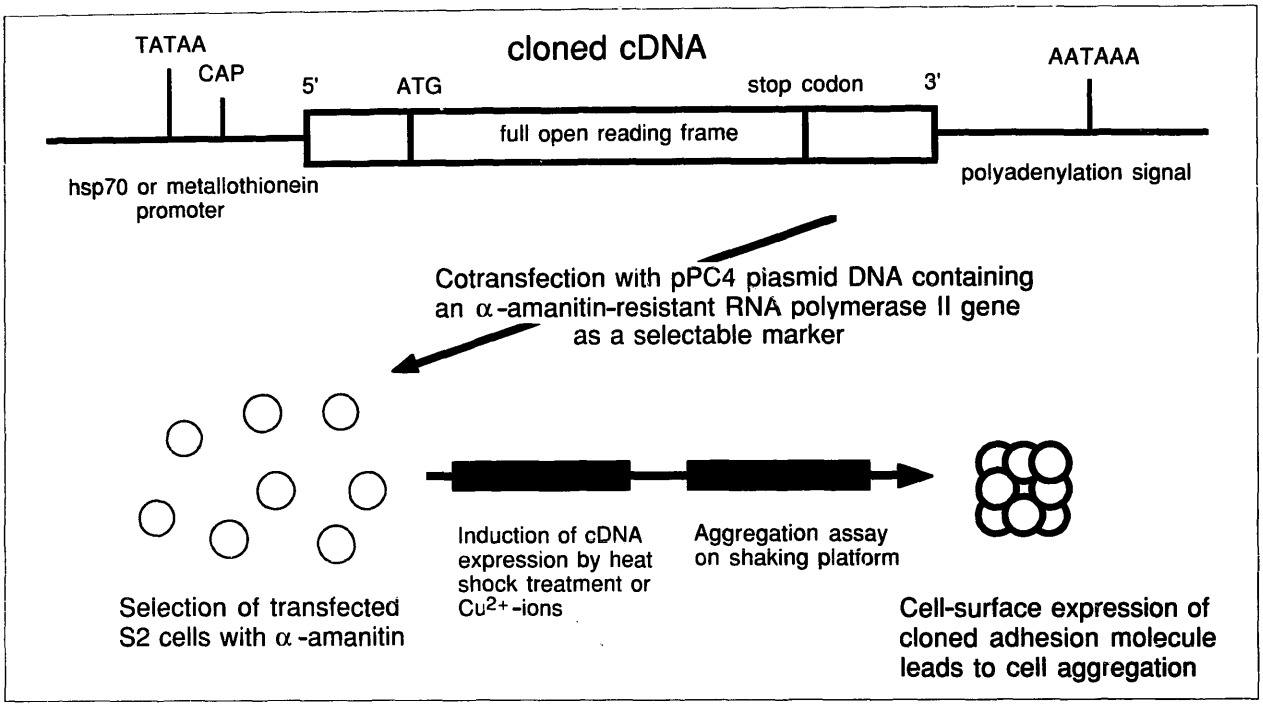

Figure 1

Expression of a cloned cDNA in S2 cells. The cDNA is first subcloned into the polylinker of a *Drosophila* expression vector (pHT4, pCaSpeR-hs or pRmHa-3). Plasmid DNA is introduced by calcium phosphate precipitation or some other transfection method into S2 cells. Cotransfection with a plasmid containing a selectable marker allows the selection of transfected cells. Expression of polypeptides encoded by the cloned cDNA is induced by a heat shock or the incubation with Cu^{2+}-ions and cell aggregation is evaluated after incubation on a shaking platform.

transfection experiments[8,9]. In addition to subcloned stable cell lines, transiently transfected cell populations have been successfully used for cell aggregation experiments. Permanently transfected, clonal cell lines can be isolated by limiting dilution or by cloning in soft agar.

Drosophila cell adhesion and recognition molecules in S2 cells

The expression of cDNAs encoding *Drosophila* cell-surface proteins in transfected S2 cells has recently been used to characterize functional aspects of these proteins. A few examples where this approach has been successfully applied will be briefly reviewed and, more specifically, different problems that were addressed using this experimental system and potential future applications to unresolved questions will be discussed.

The fasciclin I and fasciclin III glycoproteins were first identified using monoclonal antibodies that were raised against membrane-associated proteins from insect nervous systems[10,11]. Their regional expression on particular portions of embryonic axons in the developing nervous system suggested a function in growth cone guidance and the selective fasciculation (bundling) of axons. The analysis of fasciclin I and III cDNAs from *Drosophila*, however, initially revealed no sequence homologies to any known protein and therefore provided no clue to their biological function[12,13]. Fasciclin I is a glycoprotein with an apparent molecular mass of 70 kDa that consists of four separate homologous domains of 150 amino acids each. These fasciclin I repeats exhibit no homology to any known structural protein motif. The fasciclin I protein is anchored in the plasma membrane by a covalently attached glycosyl-phosphatidylinositol-lipid moiety.

At least two different protein forms are generated from the fasciclin III locus[11]. The cDNA sequence of the largest form, which has an apparent molecular mass of 80 kDa, indicates that this fasciclin III species is a transmembrane protein[12]. Its extracellular portion consists of three domains that distantly resemble the immunoglobulin protein domain, although one of these lacks the characteristic cysteine residues[14]. Expression of fasciclin I or fasciclin III protein in S2 cells established that both molecules are homophilic, Ca^{2+}-independent cell-adhesion molecules[12,15]. The homophilic nature of such cell–cell interactions can be determined by using mixtures of two different S2 cell lines, one that expresses the protein in question, while the other does not. One of the cell lines is labeled with a non-diffusible fluorescent dye, such as the lipophilic compound DiI. After the formation of cell aggregates, the cell clusters are scored for the inclusion or exclusion of DiI-positive cells. The coaggregation of both cell types indicates that the adhesion molecule in question interacts with a non-homologous ligand constitutively expressed by S2 cells, while the exclusion of the non-expressing cell line suggests a homophilic recognition mode for the protein product of the expressed cDNA.

When fasciclin I-expressing S2 cells are mixed with fasciclin III-expressing S2 cells, they undergo cell-type specific sorting into separate homogeneous aggregates[15]. Such mixing experiments using S2 cells transfected with different sets of cell-adhesion molecules will be particularly useful for the determination of cell-adhesion-molecule specificity and might reveal possible interactions between different cell-adhesion molecules. The selective and differential adhesion that plays a major role in the aggregation and sorting out of these

cells in culture might be equally important in organizing cells and tissues during embryonic development. These experiments could be a further step towards duplicating the observation *in vitro*, that dissociated vertebrate cells from different tissues can reassemble and sort into organized layers of tissues in culture[16].

The *Drosophila* fasciclin II and neuroglian polypeptides are related to two well-characterized vertebrate neural cell-adhesion molecules, NCAM and the L1 protein (also called Ng-CAM, NILE or G4-antigen) respectively[17,18]. Both are transmembrane glycoproteins with multiple immunoglobulin and fibronectin type III domains in their extracellular portions. Cell transfection experiments confirmed that these two insect proteins not only share sequence and structural similarities with these vertebrate cell-adhesion molecules but also serve a similar functional role as Ca^{2+}-independent, homophilic cell-adhesion molecules[14]. Figure 2 depicts the result of such an aggregation experiment. Whereas the parental S2 cell line or S2 cells transfected with a control construct do not adhere to each other, S2 cell lines expressing *Drosophila* neuroglian or fasciclin II protein on their surface rapidly form large cell aggregates when agitated gently on a shaking platform. The availability of such a simple functional assay system will now allow the functional dissection of these molecules by deleting or rearranging domains *in vitro*, followed by an analysis of the effects of the alterations in the cell aggregation assay. It will be of interest to determine which domains are necessary and sufficient for the homophilic adhesion phenotype of the neuroglian and fasciclin II polypeptides, and what roles the immunoglobulin protein domains and the fibronectin type III domains are playing in this process.

The *Drosophila* chaoptin polypeptide was first identified as a photoreceptor cell-specific antigen in the developing eye imaginal disc and the adult fly eye[19]. Flies that are mutant for the *chaoptin* gene exhibit a severe disorganization of the microvilli in the developing rhabdomeres and a disruption of cell–cell contacts between adjacent photoreceptor cells[20]. This suggested that chaoptin might be a cell-type-specific adhesion molecule that is necessary for proper *Drosophila* photoreceptor cell morphogenesis. The chaoptin protein is composed of 41 tandemly arranged leucine-rich repeat units and is anchored in the outer lipid leaflet of the plasma membrane by a covalently attached glycosyl-phosphatidylinositol-lipid moiety[21,22]. After introducing a cDNA encoding *Drosophila* chaoptin under the control of the hsp70 promoter into S2 cells, Krantz and Zipursky demonstrated that upon induction of the hsp70 promoter, transfected cells expressed correctly processed chaoptin polypeptide on their surface[22]. The chaoptin protein induced the formation of S2 cell clusters by a homophilic, Ca^{2+}-independent mechanism. An incubation with phosphatidylinositol-specific phospholipase C, which leads to the detachment of the protein from the cell surface, or with

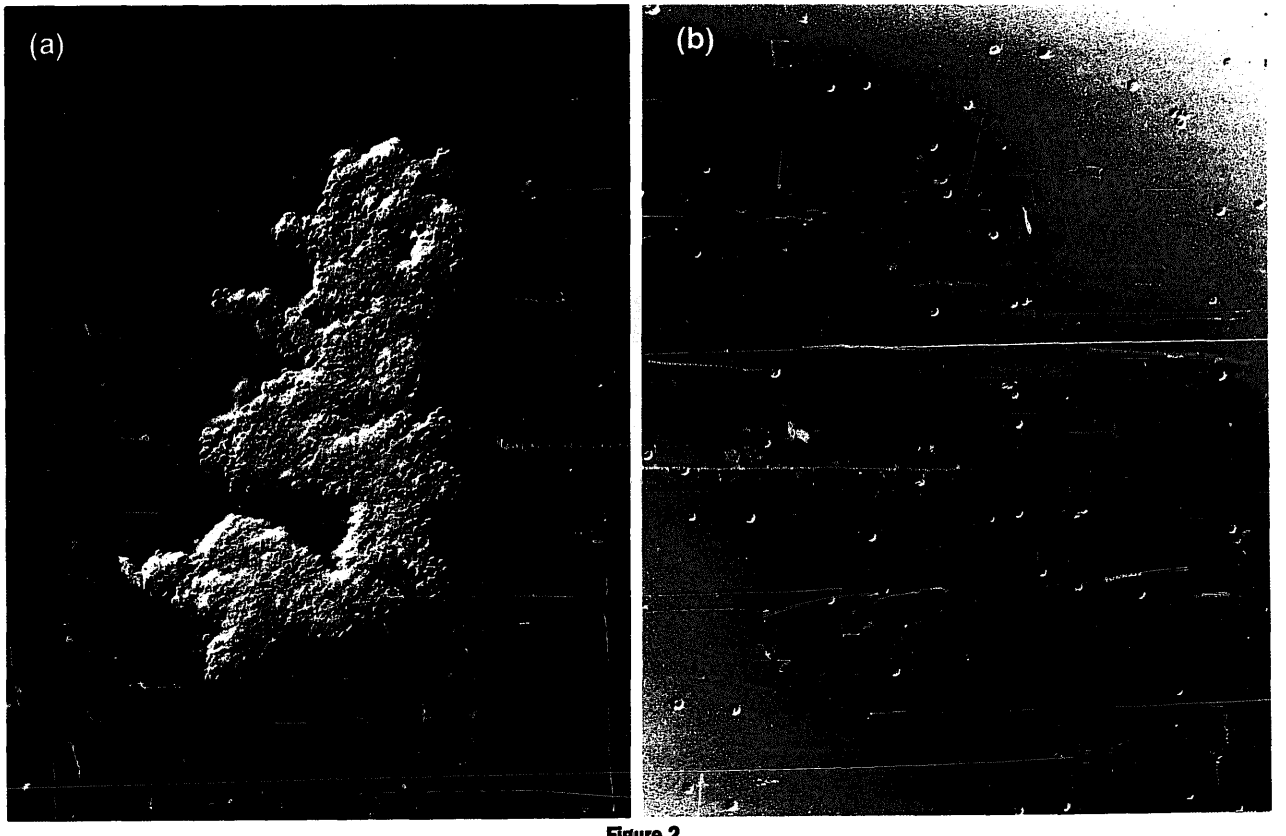

Figure 2

A functional assay for neuroglian mediated homophilic cell adhesion. A cDNA encoding the neuron specific form of *Drosophila* neuroglian was cloned into the pRmHa-3 expression vector and cotransfected into *Drosophila* S2 cells along with the plasmid pPC4, which carries an α-amanitin resistant RNA polymerase II gene from *Drosophila*. Transformants were selected by growth in 5 μg ml^{-1} α-amanitin and clonal lines were obtained by cloning in soft agar. **(a)** Cells which have been transformed with a construct in which the neuroglian cDNA is in the backwards orientation to the promoter, do not express the protein and do not undergo aggregation. **(b)** Cells which have been transformed with a cDNA in the proper orientation, express neuroglian protein upon induction with Cu^{2+}-ions and undergo a rapid aggregation response.

antibodies against *Drosophila* chaoptin interfered with the aggregation of chaoptin-expressing cells. These experiments identified chaoptin as a cell-type-specific adhesion protein.

Another molecule with leucine-rich repeats that has cell adhesion activity is the product of the *Drosophila Toll* gene[23]. The *Toll* locus is a maternal effect gene that is essential for the formation of dorsoventral polarity in the developing embryo. The Toll polypeptide is a transmembrane protein with 15 leucine-rich repeat units in its extracellular domain. Transfected S2 cells that express Toll protein after a heat shock treatment form large cell aggregates[24]. Fluorescently labeled control cells that do not express Toll protein coaggregate with Toll-expressing cells. Such mixing experiments indicate that this adhesion process is based on a heterophilic interaction between the *Toll* gene product and a receptor that is normally expressed on S2 cells. However, the nature of the Toll protein ligand is currently unknown. It could be another membrane protein or a glycolipid but it has also been suggested that leucine-rich repeats, due to their predicted amphipathic structure, can interact directly with the lipid bilayer of the contacting cell[21]. These findings suggest that other members of the leucine-rich repeat family of proteins, found in such divergent species as yeast and man, might have similar adhesive properties.

Cell transfection experiments involving the *Notch* (*N*) and the *Delta* (*Dl*) gene products provide another example for a heterophilic cell–cell interaction between two different cell-surface molecules[25]. *Notch* and *Delta* belong to a group of genes called the neurogenic genes that are responsible for the decision whether an epidermal precursor cell adopts a neuronal or an epithelial cell fate during *Drosophila* embryogenesis. It is known that cell–cell interactions between the epidermal precursor cells are of central importance for this developmental differentiation choice, and genetic experiments suggest a dose-dependent interaction between the *Notch* and the *Delta* gene products in this process. Molecular cloning of cDNAs derived from the *Notch* and the *Delta* genes indicate that both encode transmembrane glycoproteins with EGF repeats in their extracellular domains[26,27]. Notch has 36 such repeats, and Delta has 9. Fehon and co-workers introduced *Delta* or *Notch* cDNAs into S2 cells under the control of the metallothionein promotor[25]. They demonstrated that cells expressing the Notch protein bind specifically to cells expressing the Delta protein, which required Ca^{2+} ions. These results provided the first direct evidence that the *Notch* and the *Delta* gene products interact directly with one another at the molecular level. However, how the Notch–Delta protein–protein interaction influences the developmental choices taken by the cells involved remains unknown.

The analysis of heterophilic cell-adhesion reactions is of specific interest, since such interactions might provide ways of sending signals unidirectionally from cell to cell or of delivering differentiation cues, as in the developing neuroectoderm. However, unless both partners in a heterophilic interaction are known, finding an unknown receptor counterpart still presents a significant task. As illustrated by the *Drosophila* neurotactin molecule, exogenous expression of a cloned receptor cDNA in S2 cells might provide a valuable tool to approach this problem. Neurotactin represents a transmembrane glycoprotein with an apparent molecular mass of 135 kDa[28–30]. Its extracellular domain shows a strong structural homology to serine esterases without retaining the active site of this group of enzymes. It is initially widely expressed in the *Drosophila* embryo and later becomes restricted to parts of the nervous system, especially at points of cell–cell contacts. However, S2 cells expressing the neurotactin protein on their surface are completely inert in the kind of cell aggregation experiment described above[30]. This excludes the possibility that the neurotactin molecule by itself can mediate homophilic cell–cell binding. However, cells that are transfected with a neurotactin cDNA bind to a subpopulation of primary embryonic cells that apparently express a ligand to the neurotactin molecule on their surface. Since the nature of this potential neurotactin ligand is currently unknown, the availability of this cellular binding assay will hopefully lead to its identification.

The expression of vertebrate cell-adhesion molecules in tissue culture cells

The transfection of cloned cDNAs encoding potential cell-adhesion molecules into tissue culture cells was first introduced by Takeichi and co-workers for members of the cadherin family of vertebrate Ca^{2+}-dependent cell-adhesion proteins[1,31,32] and by Edelman and co-workers for other vertebrate cell adhesion molecules[2,33,34]. These groups transfected cDNAs coding for different cadherin species or different forms of the neural cell-adhesion molecule (N-CAM) into mouse liver (L) cells[1,2,31]. Constitutive expression of cadherin or N-CAM cDNAs was driven by the SV40 early or a composite SV40/*Herpes simplex* virus thymidine kinase gene promoter. The mouse sarcoma 180 or Neuro 2a cells have been also lately used as hosts for the expression of different vertebrate cell-adhesion molecules, such as liver cell-adhesion molecule (L-CAM/E-cadherin), N-cadherin and N-CAM[32–34]. Compared to the S2 cell system all these other cell transfection systems have some limits or shortcomings. Aggregation and sorting assays in suspension are usually performed after a trypsin pretreatment to resuspend and dissociate the cells[1,2]. This precludes the investigation of cell-adhesion proteins that are especially susceptible to proteolysis. The expression vectors that were used in these studies are based on viral promoters and do not permit the regulation of cDNA expression. Secondary protein modifications such as the addition of fatty acids, specific oligosaccharides or phosphate groups might be important regulators of cell adhesion and cell–cell interactions *in vivo*. *In vitro* cell expression systems (such as the one discussed in this review) might be incapable of performing all these modifications and therefore could be of limited use for the investigation of such biologically relevant functional modulations. Mouse L cells, for example, fail to anchor the smallest N-CAM form with a glycosyl-phosphatidylinositol-lipid in the plasma membrane and secreted the newly synthesized molecules instead[2]. In contrast, S2 cells are capable of performing this secondary protein modification correctly as has been shown for fasciclin I and chaoptin, two glycosyl-phosphatidylinositol-lipid anchored molecules[15,22]. The use of the *Drosophila* S2 cell line might also provide a good alternative for people working with vertebrate cell-adhesion molecules. The low intrinsic adhesiveness of S2 cells makes them an ideal host cell line for any type of 'sticky' molecules.

Acknowledgements

We thank C. Goodman, M. Seeger and M. Burmeister for comments on the manuscript and A. Nose for useful discussions.

References

1 Nagafuchi, A. *et al.* (1987) *Nature* 329, 341–343
2 Edelman, G. M. *et al.* (1987) *Proc. Natl Acad. Sci. USA* 84, 8502–8506
3 Schneider, I. (1972) *J. Embryol. Exp. Morph.* 27, 353–365
4 Schneuwly, S., Klemenz, R. and Gehring, W. J. (1987) *Nature* 339, 61–64
5 Rubin, G. M. and Spradling, A. C. (1983) *Nucleic Acids Res.* 11, 6341–6351
6 Bunch, T. A., Grinblat, Y. and Goldstein, L. S. B. (1988) *Nucleic Acids Res.* 16, 1043–1061
7 Jockerst, R. S., Weeks, J. R., Zehring, W. A. and Greenleaf, A. L. (1989) *Mol. Gen. Genet.* 215, 266–275
8 Bourouis, M. and Jarry, B. (1983) *EMBO J.* 2, 1099–1104
9 Rio, D. C. and Rubin, G. M. (1985) *Mol. Cell. Biol.* 5, 1833–1838
10 Bastiani, M. J., Harrelson, A. L., Snow, P. M. and Goodman, C. S. (1987) *Cell* 48, 745–755
11 Patel, N. H., Snow, P. M. and Goodman, C. S. (1987) *Cell* 48, 975–988
12 Snow, P. M., Bieber, A. J. and Goodman, C. S. (1989) *Cell* 59, 313–323
13 Zinn, K., McAllister, L. and Goodman, C. S. (1988) *Cell* 53, 577–587
14 Grenningloh, G. *et al. Cold Spring Harbor Symp. Quant. Biol.* (in press)
15 Elkins, T. *et al.* (1990) *J. Cell Biol.* 110, 1825–1832
16 Steinberg, M. (1963) *Science* 141, 401–408
17 Bieber, A. J. *et al.* (1989) *Cell* 59, 447–460
18 Harrelson, A. L. and Goodman, C. S. (1988) *Science* 242, 700–708
19 Fujita, S. C. *et al.* (1982) *Proc. Natl Acad. Sci. USA* 79, 7929–7933
20 Van Vactor, D., Jr, Krantz, D. E., Reinke, R. and Zipursky, S. L. (1988) *Cell* 52, 281–290
21 Reinke, R., Krantz, D. E., Yen, D. and Zipursky, S. L. (1988) *Cell* 52, 291–301
22 Krantz, D. E. and Zipursky, L. S. (1990) *EMBO J.* 9, 1969–1977
23 Hashimoto, C., Hudson, K. L. and Anderson, K. V. (1988) *Cell* 52, 269–279
24 Keith, F. J. and Gay, N. J. (1990) *EMBO J.* 9, 4299–4306
25 Fehon, R. G. *et al.* (1990) *Cell* 61, 523–534
26 Wharton, K. A., Johansen, K. M., Xu, T. and Artavanis-Tsakonas, S. (1985) *Cell* 43, 567–581
27 Vässin, H., Bremer, K. A., Knust, E. and Campos-Ortega, J. (1987) *EMBO J.* 6, 3431–3440
28 Hortsch, M. *et al.* (1990) *Development* 110, 1327–1340
29 De la Escalera, S. *et al.* (1990) *EMBO J.* 9, 3593–3601
30 Barthalay, Y. *et al.* (1990) *EMBO J.* 9, 3603–3609
31 Hatta, K., Nose, A., Nagafuchi, A. and Takeichi, M. (1988) *J. Cell Biol.* 106, 873–881
32 Matsunaga, M., Hatta, K., Nagafuchi, A. and Takeichi, M. (1988) *Nature* 334, 62–64
33 Mege, R-M. *et al.* (1988) *Proc. Natl Acad. Sci. USA* 85, 7274–7278
34 Friedlander, D. R., Mege, R-M., Cunningham, B. A. and Edelman, G. M. (1989) *Proc. Natl Acad. Sci. USA* 86, 7043–7047

LETTERS

Tracing origins with molecular sequences: rooting the universal tree of life

In his recent *TIBS* article[1] Lake correctly pointed out the importance of molecular sequences in determining the phylogenetic relationships among diverse groups of organisms. The main emphasis of his article was on results that were found to be unreliable either because of insufficient data or because of possible artifacts stemming from unequal rates of evolution. I fear that many readers will get the impression that molecular data used as evolutionary markers are a very unreliable tool.

All algorithms available for phylogenetic analyses rely on assumptions that are sometimes difficult to verify but artifacts and unreliable results can usually be easily detected if the investigator takes care to compare the outcome of different algorithms and alignments. In many cases, distance matrix, maximum likelihood, parsimony and evolutionary parsimony analyses do give the same results (see, for example, Ref. 2). Thus, although the individual algorithms tend to overestimate the reliability of the results, and ignore a possible bias due to a particular alignment or algorithm, taken together they lend credibility to the obtained results.

A good case to illustrate this point is the rooting of the universal tree of life by means of genes that had already undergone a gene duplication in the last common ancestor. By use of the DNA sequences encoding the catalytic (i.e. ATP hydrolysing) and non-catalytic (i.e. ATP binding, but not hydrolysing) subunits of F-, V- (vacuolar) and archaebacterial proton pumping ATPases it was shown that these ATPases are homologous to each other[2,3]. In addition, analysis showed that the gene duplication that gave rise to the catalytic and non-catalytic subunits occurred before the lines that lead to the three Urkingdoms or domains separated from each other[2]. Thus one can use the non-catalytic subunits as an outgroup to root a tree which uses the catalytic subunits as markers for the organismal evolution (see Fig 1). Use of the non-catalytic subunit as an organismal marker and the catalytic subunit as an outgroup gives the same result.

Use of ATPase subunits as evolutionary markers suggests that the archaebacteria branch off from the line that leads from the last common ancestor to the eukaryotes. This result has also been obtained for *Sulfolobus*[1], *Methanococcus*[4], *Methanosarcina* (H. Kibak, J. P. Gogarten and L. Taiz, unpublished; sequence from

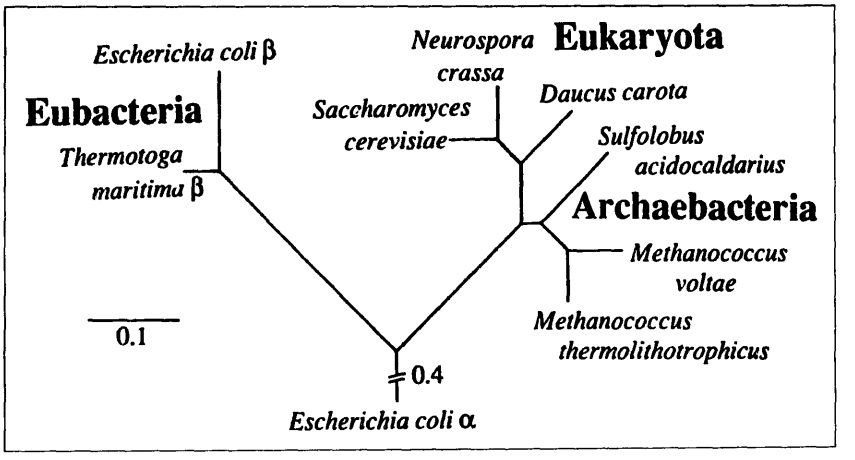

Figure 1

Phylogenetic tree showing the relations between the three Urkingdoms (domains). The topology and the branch lengths were calculated using Felsenstein's maximum likelihood method[14]. Branches are scaled in terms of the probability for change of the first base of the codon. Parameters for the algorithm, sequences and their alignment were as described in Ref. 4. All branch lengths were calculated to be positive at the 1% significance level. Using evolutionary parsimony the archaebacterial tree was significantly supported with $p<2\%$.