What do we mean by “cell type”? Two articles in this issue of Circulation Research address this question in different ways. The first study, by Dube et al,1 uses very well-defined tools to define endothelial differentiation at the level of transcriptional control. The second study, by Kowal et al,2 looks for novel genes that distinguish two cell types. These studies, taken together, illustrate a major change in how we define cell type, a change that is about to be accelerated by the power of systematic genomics.

Dube et al1 use in vitro systems to explore the promoter structure for the endothelial cell–specific receptor tyrosine kinase Tie2. Tie2 is a receptor for both angiopoietin-1 and angiopoietin-2. Like vascular endothelial growth factor, angiopoietin-1 is essential for normal vascular development whereas angiopoietin-2 is a naturally occurring antagonist for angiotensin I and Tie2. Thus, regulation of expression of the Tie2 receptor is likely a key issue in the formation of blood vessels.3,4 Dube et al1 identify a putative Ets transcription site in the Tie2 promoter. One of the proteins binding to this site, NERF2, can be shown to promote transcription in endothelial cells but not in the other cells studied. Related NERFs and other Ets binding factors either did not have this activity or were not found in the cultured cells used by these authors. Although mechanistically solid for transfected endothelial cell lines in vitro, the studies presented by Dube et al do not provide evidence that NERF2 is expressed in endothelial cells in vivo, and the in vitro studies only use Chinese hamster ovary, 293, and CV-1 cell lines as comparison cells. These aneuploid, heterogeneously derived cell lines may or may not provide a representative comparison with the endothelial cell strains used in the present study, much less with true nonendothelial cell types. As NERF2 is explored, we will inevitably see studies on the in vivo expression pattern of NERF2, knockout and overexpression studies, analysis of the NERF2 promoter, and so on.

The study by Kowal et al2 takes an entirely different approach, an approach we might call “systematic exploration.” The study uses a subtractive polymerase chain reaction method to collect the set of genes enriched in one type of cell versus another. The underlying hypothesis in this study is that the cell type we call “smooth muscle cells” is actually a complex mixture of several different cell types or subtypes. A variety of evidence, much of it from our laboratory, has supported this hypothesis.5 We know that cultured smooth muscle cells obtained in different ways have different stable phenotypes.5,6 Approximately 100 genes have been shown to have differential expression between smooth muscle cells in different states and different locations. Because many of the distinctions based on cell strains studied in culture also appear in the pattern of expression of neointimal cells formed after injury, we have suggested that these different smooth muscle cell types in culture somehow represent a model for the cells forming the neointima in vivo.5 The possible existence of more than one smooth muscle subtype is especially important because we know that the atherosclerotic plaque arises from a clone of arterial smooth muscle and that smooth muscle cells can arise not only from the traditional sources but from hematogenic stem cells or neural crest.7,8

The identity of subtypes, if they exist for smooth muscle, could be as important for vascular biology as the identification of lymphocyte subtypes has been for immunology. Systematic analyses of patterns of gene expression should allow us to define these putative cell types. As a result of the genome project, we all expect in the near future that an inexpensive array will contain representative, quantitative samples of each gene.9 Analyses with these arrays will, in one fell swoop, define the spectrum of gene expression that defines two cell types or the response of a cell to some specific stimulus. The Kowal et al2 study does not use arrays. However, the subtraction-suppression hybridization method, based on the original representation distribution analysis method used for DNA,10 is theoretically an exhaustive method that may allow one to amplify the complete set of sequences that are differentially expressed in one RNA set versus another. The modifications used in the PCR Select kit (Clontech) used by Kowal et al2 include an extensive hybridization of “target” versus “driver” sequences before the amplification steps. This step should normalize the cDNA representations, leaving only sequences unique or enriched in the “target” population to be amplified. The end result of this procedure is a small tube containing a representation of all the differentially expressed sequences contained between a frequent cutter. With enough money, it is possible to completely sequence the content of the PCR Select subtractive product and, by doing so, one should be able to collect a definitive set of the differentially expressed genes that represent that cell type.

One implication of definitive analyses of gene expression should be a redefinition of what we mean by cell type. Currently “cell type” is defined, in most cases, either by morphology or by a small number of genes that have been identified with that tissue. Thus, we define liver either by its morphology or by its expression of albumin. We define endothelium first by its morphology and then by looking for all type-specific factors, such as the NERF2 transacting factor described by Dube et al.1 If one were to discover, however, that a cell with the properties...
of a renal tubular epithelial cell also made albumin, then our definition of hepatocyte would no longer be tenable. Definitive analyses of gene expression will either give us more confidence in our ability to identify cell types or, unfortunately, undermine our current clear view of cell types.

A critical issue in these sorts of systematic examinations of expression pattern will be the choice of cell lines used as prototypes. The PAC1 cell line and the cardiac cells in the Kowal et al study do not bear a clear relevance to smooth muscle cell types. We already know that PAC1 is aneuploid and only partially similar to in vivo smooth muscle (A.B. Firulli, D.K. Han, L. Kelly-Roloff, S.M. Schwartz, E.N. Olson, J.M. Miano, unpublished data, 1999). Thus, while EVEC is interesting as a chance finding, the choice of cells for the original analysis cannot provide any systematic data on smooth muscle cell types. A better example of systematic analysis as a basis for exploring cell identity may be our identification, in collaboration with Drs Miano and Olsen from the Kowal et al study, of homeotic genes differentially expressed by human fetal versus adult smooth muscle.11

The challenge of systematic gene exploration will not be discovering genes, but fitting them into functional pathways.12,13 Comprehensive views of gene expression are already a reality for the simple expressed genome of the yeast where the use of arrays to define clustered patterns of expression is offering new insights into the mechanisms involved in the cell cycle or sporulation.12 Similar studies have recently been published looking for clustered expression patterns in the response of mammalian cells to serum or lymphocyte activation.14,15 Several sets of mammalian genes clustered by expression pattern in the response to serum in vitro are already known as early response genes and intermediate response genes.15 Many of these are also represented in the neointima formed when an artery responds to balloon injury.5

One would hope that cluster analysis would lead to novel insights by identifying groups of proteins important to a process, such as intimal formation. One example of such a cluster may come from a recent study done in our laboratory. In an effort to add to the definition of our two smooth muscle subsets, we used a mini-array consisting of 40 random clone cDNAs derived from these cells. Of these 40, we found 5 that differentially and consistently mark the “pup” phenotype as opposed to the adult phenotype. To our surprise, one of these genes was ZO2, a tight junction protein. Smooth muscle cells are not supposed to have tight junctions. Further exploration, however, showed us that other members of the tight junction-associated complex are associated with the pup phenotype as well as with the formation of the neointima. In addition, application of subtraction-suppression hybridization to the pup-adult system has identified one new transcript that appears to interact with the ZO2 complex (T. Ito, S.M. Schwartz, unpublished data, 1999). Although smooth muscle cells do not form canonical tight junctions, the morphogenic functions of this set of proteins are likely to be important in the formation of the intima (L.D. Adams, J.M. Lemire, S.M. Schwartz, unpublished data, 1999).

The identification of EVEC may merit a similar systematic exploration of functionally related genes. EVEC is a member of the elastin microfibril superfamily. The first gene to be identified as differentially expressed in the vascular response to injury was elastin, and overexpression of elastin and fragmentation of elastin were already known to be implicated in the atherosclerotic plaque.5 Elastin itself is related to the neural crest origin of a subset of smooth muscle cells.16 Mutations in elastin are responsible for supravalvular aortic stenosis,17 and Marfan syndrome is attributable to mutations in fibrillin, the first elastin microfibril to be defined.18 It is intriguing to wonder if the identification of EVEC as an intimately expressed gene is “telling us” that a more systematic understanding of elastin would be valuable.

In summary, biology is changing. We have the tools to explore hypothesis-based studies in increasingly minute detail. At the same time, systematic exploration will provide us with more pathways to be defined. As such studies proceed, the functional map will get more and more complex. The challenge will be to discriminate critical mechanistic steps from less important details.

References

Key Words: array display • cell lineage • differentiation • cell type

The Definition of Cell Type
Stephen M. Schwartz

Circ Res. 1999;84:1234-1235
doi: 10.1161/01.RES.84.10.1234

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/84/10/1234