

TECHSIGHTING
MOLECULAR MARKERS
 Mailbox:
www.sciencemag.org/dmail.cgi?53802a

Tissue Cork Borer

Richard Peters and Robert Sikorski

Finding new tools for the diagnosis, prognosis, and treatment of cancers is a major goal of cancer research. Tumor markers are particularly useful for both cancer diagnostics and prognosis. New tumor markers based on genetic polymorphisms are routinely identified. The analysis of these markers requires a variety of technologies, from immunofluorescence, to in situ hybridization, to the polymerase chain reaction.

Unfortunately, to date, tumor markers have not gained widespread use in day-to-day clinical practices, although they are used in certain clinical trials to guide therapy. One of the main reasons for this is the paucity of solid population data to show the predictive value of any one marker. Additionally, it is difficult to look at an array of tumor markers simultaneously in single specimens. A multitest survey of tumors is needed to show the predictive relationships among different markers.

A recent report in *Nature Medicine* may well change the medical landscape and speed implementation of such markers in clinical practice. The authors describe a new technique that allows researchers to study up to 1000 tumor biopsies at a time (1). The authors built an instrument, which consists of a thin-walled, stainless steel tube, with an inner diameter of 600 μm , sharpened like a cork borer. They used this needle to select punch biopsies 3 to 4 mm in height from fixed tumor samples. Using a solid stainless steel wire, which functions as a stylet, the sample is then emptied into a recipient array block with drilled holes. A digital micrometer moves the system to successive locations in the growing array. The recipient paraffin block (45 mm by 20 mm) can have 200 consecutive 8- μm sections [see figure 1 of (1)]. Each of the 200 sections holds up to 1000 tumor samples. The power of this tissue microarray technique is the capability of performing a series of analyses of 1000 specimens in a parallel fashion. Armed with such a research tool, researchers can study vast numbers of tumor samples in a short time and can generate a wealth of data on the application of tumor markers.

As proof of principle, the authors performed immunohistochemical analyses (for protein expression), fluorescence in situ hybridization (for DNA amplification), and RNA/RNA in situ hybridization (for messenger RNA expression). They used

645 breast tumors to create the microarray. They then studied sections of this array with eight markers: estrogen receptor expression, p53 expression, and DNA amplification at the CCND1, ERBB2, MYC, 17q23, 20q13, and MYBL2 loci. Their results match published data, validating their approach. Specifically, at least one of the six DNA loci was amplified in 52% of primary tumors: ERBB2, 18%; MYC, 25%; CCND1, 24%; 17q23, 13%; 20q13, 6%; and MYBL2, 7%. They also reproduced previous findings showing that ERBB2 and MYC were amplified more often in tumors lacking estrogen receptors, whereas the opposite was observed for CCND1.

Finally, p53-positive tumors had a higher frequency of MYC and ERBB2 amplification than p53-negative ones. The technique appears simple, and the authors can process 1000 samples in 3 days time. The major drawback of the approach is the small diameter (0.6 mm) of the punch biopsy. With such a small core, areas of interest are easily missed in tumors that show a significant amount of heterogeneity. Multiple samples from each tumor specimen can reduce this problem. The authors note that increasing the density of the array so that more than 1000 specimens can be housed in the recipient block would extend the throughput of the system. Other technical tricks could be implemented; for instance, researchers could combine immunohistochemical methods with fluorescence in situ hybridization to use multiple detection probes simultaneously.

Variations and improvements on this technique can be expected in the near future. This research tool should have wide application in the field of cancer research, as well as in other fields (such as developmental biology) where molecular markers are especially informative.

References

1. J. Kononen et al., *Nature Med.* **4**, 844 (1998).



A Frankenstein Experiment

Robert Sikorski and Richard Peters

Whenver the process of exchanging reagents among scientists becomes easier, science benefits. Take as an example the community of scientists who experiment with the simple organism, yeast. The simple property that

yeast strains can be dried on filter paper disks has aided the distribution of a wide variety of mutants. Dried yeast strains can be maintained indefinitely at room temperature. They can be transported in an envelope for just 32 cents. Imagine, though, if yeast strains needed to be cryopreserved in liquid nitrogen to maintain their viability. The cost alone of exchanging the thousands of mutant yeast strains would be enormous.

Another community of researchers, those studying the mouse, is also trying to manage the ever-increasing numbers of their mutant strains. Mouse strains are usually passed between laboratories in the form of viable animals. Mouse embryos can be frozen, but the process is technically challenging to do in a reproducible way. Also, it is a drain on resources to go through the process of making and saving "libraries" of frozen embryos, merely for distribution. Maintaining a working colony of mice is expensive and time-consuming. Maintaining extra mice for colleagues' requests now adds to the lab budget.

A possible solution for the mouse community, and perhaps others, is in a recent issue of *Nature Biotechnology*. A group from Hawaii has succeeded in bringing freeze-dried mouse sperm back to life (1). There have been numerous attempts to develop techniques for low-cost sperm preservation, but they have all yielded the same failing results. At the end of these procedures, the sperm are essentially dead: They do not move and have no plasma membranes.

However, the researchers—led by Ryuzo Yanagimachi—tried a different approach. First, they created vials of freeze-dried sperm samples by freezing sperm in liquid nitrogen and drying them into powder form under a vacuum. No special buffers or technology was used. They then rehydrated the dried sperm and examined them for viability. As before, all sperm looked dead by conventional tests, such as special fluorescence stains.

Next, they performed an experiment a bit like Frankenstein's: They used micromanipulation to remove the freeze-dried sperm heads and to inject them directly into unfertilized mouse oocytes. The idea was to see whether the nucleus was dormant and could be revived within the cytoplasm of a host egg. With this crude process, they achieved rates of fertilization >90%. In the end, 30% of all head-injected oocytes produced viable mouse offspring that appeared completely normal.

The straightforward techniques used by these researchers, micromanipulation and vacuum drying, should be applicable in any major mouse research laboratory. It remains to be seen whether the sperm of other species will be able to withstand

freeze-drying. If so, this would lower the storage and shipping costs of animal strains considerably.

References

1. T. Wakayara and R. Yanagimachi, *Nature Biotechnol.* **16**, 639 (1998).

NET TIPS

META-LANGUAGE

Mailbox:

www.sciencemag.org/dmail.cgi?53804

XML Is Hatching

Robert Sikorski and Richard Peters

The intense promotion of Internet technologies makes it difficult for the busy scientist to separate fact from fiction. At the cutting edge, it is often pointless to try, since most new ideas fail to gain widespread acceptance in the marketplace. It is often best to watch as embryonic technologies mature in their "larval" stage, looking for signs that the technology is being incorporated into current software and tools. Use in solving practical problems and adoption by major organizations and companies are also good signs. However, many Internet "inventions" never make it to these last stages.

One technology that has now developed to the point of hatching and is well on its way to leaving the nest is the extensible markup language, or XML. XML is being developed by a large and dynamic group under the leadership of the W3 consortium, the group that brought you HTML, the language of the World Wide Web.

In the simplest sense, XML can be viewed as a way of attaching additional layers of meaning to any word or words. This is done by incorporating meta-information into the text in the form of tags. A text selection with XML tags will have the following structure: `<TAG>Item</TAG>`. The tags can be nested to create complex organizations. The additional meaning, or meta-data, that the tags deliver can be extracted easily by software tools. Simply viewing the tags by eye is useful. In an XML document you can extract selected pieces of data from the body of the text. If this all seems too abstract, a concrete example follows.

Let's take the field of human genetics, where gene mutation data are a staple product of this research. Unfortunately, the actual data in most genetics manuscripts (the mutation positions, the phenotypes, the assays used, and so forth) are all buried in the body of the text at publication time. For instance, to review the world's literature for all of the mutations in the gene *p16* that cause melanoma, you would have to find the hundreds of papers that mention *p16*, read them

all, extract the data manually, correct the data for the different numbering schemes used in several hundred different mutations, and so on. The task is daunting, and it underscores the need for a different way to handle this type of scientific data. Mutation databases are being formed. Such databases are easy to start, but they are very difficult to maintain in the long run. Given the numbers of genes studied, it is hard to see how high-quality databases can be kept for each one. Also, without standards, it is difficult to compare or collate data from different databases. Instead, what if we just assigned a few XML tags to the published data as it resides on the publisher's Web site? Add tags such as `<Gene>p16</Gene>` or `<Mutation>E30A</Mutation>` and it would now be easy for software tools to extract the actual data automatically, collect it into one data structure, and allow you to visualize it in many different ways. You could feed many papers into an XML grinder to get data from multiple sources. All you need for this to work is a common set of tags to describe mutations. This set does not yet exist for genetics, but it is easy to imagine by just extrapolating from work done in other fields.

In the field of chemistry, an advanced set of XML tags has been created and deployed already. Termed Chemical Markup Language (CML), the tags provide a way of describing the structure of any molecule—the atoms, the bonds, the isotopic constitution. Think of CML as a networked version of standard chemical nomenclature.

XML is now emerging as the standard method for passing structured data over the Internet. All major Internet-based companies have supported XML in this regard. In fact, Microsoft has embedded XML into the next version of its popular Office product. With this new release, Microsoft Word documents will be automatically translated into XML. This alone means that the general use of XML documents will rise considerably.

The software development community as a whole is now focusing on incorporating XML into everything from browsers to databases. Currently, the 4.0 version of Microsoft Internet Explorer can process and display XML documents. The 5.0 versions from both Microsoft and Netscape will handle XML as well. This will make XML the de facto standard for distributed, tagged data. Additionally, many tools for creating XML and interacting with existing databases are in early releases today. Over the next few months, there will be many to choose from.

Given the large support for XML and its incorporation into the infrastructure of the next wave of Internet computing, the field of science will realize many benefits from this technology.

SITEFINDER

Robert Sikorski and Richard Peters

XML is poised to revolutionize how scientific data is exchanged and used. Below is a collection of Web sites on XML useful to both experts and beginners. Links to these sites, as well as additional XML resources, can be found at: www.medsitenavigator.com/tips

Introduction

www.byte.com/art/9801/sec5/art18.htm

www.byte.com/art/9803/sec5/art5.htm

www.forbes.com/asp/redir.asp?/tool/html/98/apr/0413/side1.htm

Background

www.w3.org/XML

XML FAQs

<http://builder.cnet.com/Authoring/Xml20/index.html>

www.ucc.ie/xml

www.textuality.com/xml/faq.html

www.isgmlug.org/n3-1/n3-1-18.htm

CML FAQs

<http://sun01.iigb.na.cnr.it/omf/cml-1.0/doc/faq/index.html>

News

www.arbortext.com/xmlnews.html

Case Studies

www.microsoft.com/xml/scenario/chrysal.asp

XML/SGML Comparisons

www.sil.org/sgml/sgml.html

Company White Papers

Microsoft

www.microsoft.com/xml/default.asp

Sun

<http://sunsite.unc.edu/pub/sun-info/standards/xml/why/xmlapps.htm>

IBM

www.infoworld.com/cgi-bin/displayStory.pl?/features/980706xml.htm

XML Specifications

www.w3.org/TR/

WD-xml-lang-970331.html

Tech.Sight is published in the third issue of each month. Contributing editors: Robert Sikorski and Richard Peters, Medsite Communications, Boston, MA 01945, USA. Send your comments by e-mail to techsight@aaas.org or via the Web to the mailbox URL with each item.