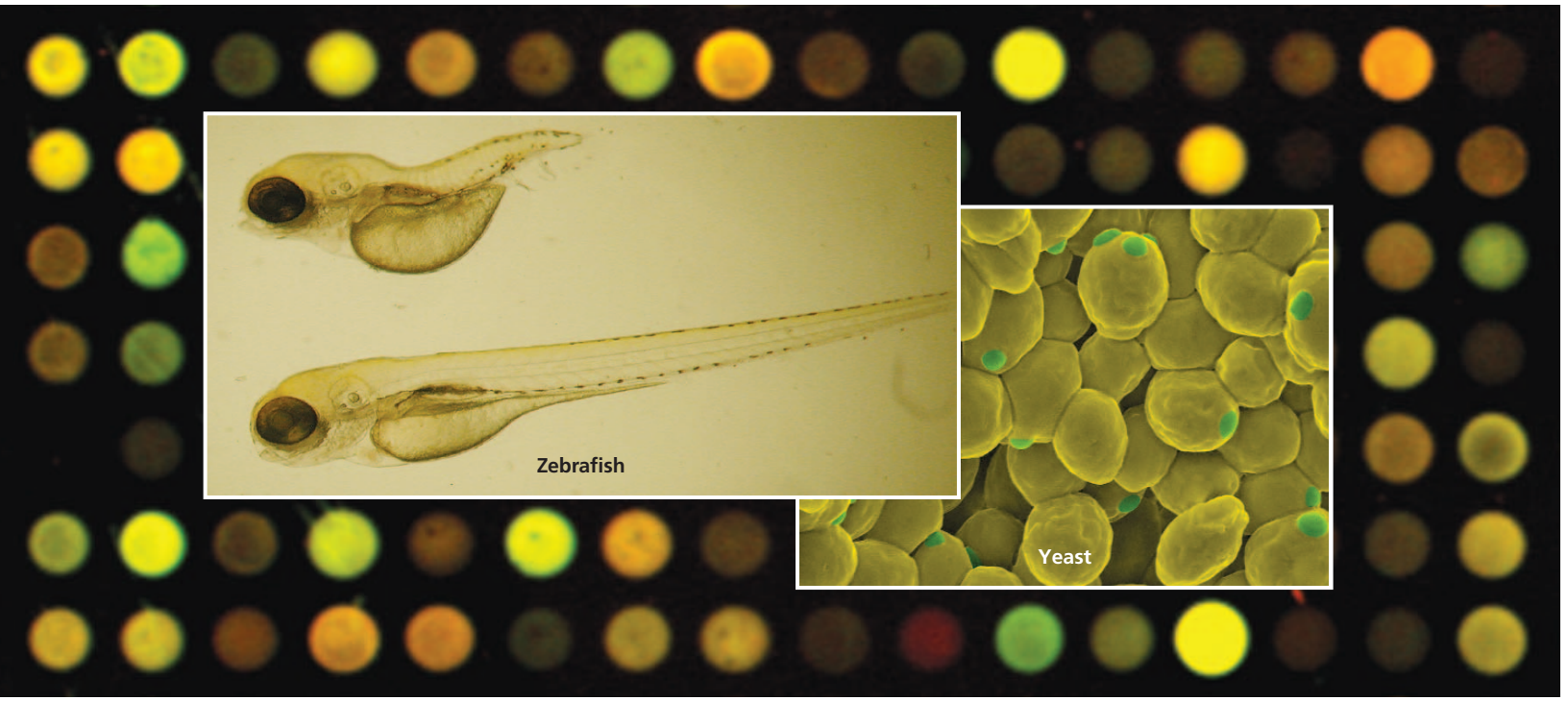
Drawing Comparisons at Duke

Comparative genomics is the study of genomes from different species in order to better understand how species have evolved and what the functions are of both genes and noncoding sections of the genome. As part of the NIEHS Toxicogenomics Research Consortium (TRC), researchers at the Duke Center for Environmental Genomics are focusing on comparative genomics as a way to explore how environmental stresses affect human health.



The NIEHS established the TRC in November 2001 to serve as the extramural mechanism by which the NCT applies microarray technology. Funded at $37 million over five years, the TRC is a coordinated, multidisciplinary effort between the NIEHS Microarray Center and five academic research institutions (Duke, the University of North Carolina–Chapel Hill, the Massachusetts Institute of Technology, Oregon Health & Science University, and the Fred Hutchinson Cancer Research Center/University of Washington), along with private companies contracted to assist with microarray development and bioinformatics. The TRC is intended to significantly advance the pace of discovery in toxicogenomics, while simultaneously accelerating the development and validation of microarray technology.

At roughly the halfway point of the initial grant period, the TRC is already bearing important scientific fruit and appears poised to make substantial contributions to the field as research supported by rapidly progressing technology produces new insights into the complex relationship between environmental exposures and gene expression.

All the Organisms, Two by Two . . .

The Duke investigators believe that comparative genomics—the isolation and identification of common, conserved genomic responses across different model species—will have a major impact on advancing useful knowledge within the field of toxicogenomics. The discipline makes it possible to probe into the phylogenetic origins of gene families and how they have been altered as species rose higher; eventually these gene structures and relative functions can be compared to those of humans, says NCT deputy director James K. Selkirk. Part of the TRC's mission has been the development of sophisticated, reliable microarrays of the genomes of model systems such as *Caenorhabditis elegans*, zebrafish, yeast, and mouse, which allow rapid, high-throughput screening across species in a variety of areas of interest.

Ongoing independent exploration by Elwood Linney and Marcy Speer into the genetics of neural tube development and defects is a case in point. Microbiologist Linney is working with zebrafish microarrays to identify sets of genes that might be involved in embryonic neural tube development, or disruption of normal development by environmental insults. Those candidate genes can then be used by human geneticist Speer to screen within subject families for polymorphisms that might correlate with families with neural tube defects such as spina bifida.

Although they have already made substantial strides in the biology, Linney says, "a major part of the progress to date has been just the technology. When we started, there weren't any microarrays for zebrafish." With the recent delivery of a 22,000-gene zebrafish microarray developed in cooperation with TRC contractors, "we're ready to start generating a variety of different types of data using perturbations of neural tube development," he says.

Such large-scale genomic screening lends itself to ferreting out the signaling pathways that might be involved in toxic response and repair mechanisms. Microarray technology allows researchers to cast a big net over a problem, rather than focusing on specific genes, says Linney. Once patterns of expression are identified via microarray, RNA interference (RNAi) or morpholinos (another gene silencer) can be applied to confirm the function of the genes. "If we think a certain toxicant is

Background: David Schwartz; insets, l–r: David Schwartz, Dennis Kunkel Microscopy, Inc.

affecting a certain gene product, we can test that by designing something to knock down that gene product and see if we still get the same phenotype," he says.

TRC principal investigator David Schwartz is using a similar comparative genomics approach in his TRC independent project. Schwartz is examining the genes and genomic responses involved in the immune response to bacterial toxins such as endotoxin, which are released into the bloodstream during bacterial infection and can, in themselves, cause a variety of symptoms. "We are using genomics as a way of highlighting a number of genes that we know are biologically related to the immune response to bacterial toxins," he explains. This will help identify genes that may have variants, some of which would predispose individuals to experience adverse responses when they have various types of infection.

Schwartz's group has identified a few very promising candidates as a result of studies with mouse microarrays. "We've found specific areas of the genome that are clearly associated with the biologic response," he says, "and we've found a couple that we think may be critical in terms of regulating the response—genes that had not previously been described as being important or relevant to processing or responding to bacterial toxins." To further elucidate the potential role of these genes, Schwartz's team is currently testing them in *C. elegans*, as well as looking at loss of function by using RNAi to downregulate their expression in mammalian cells.

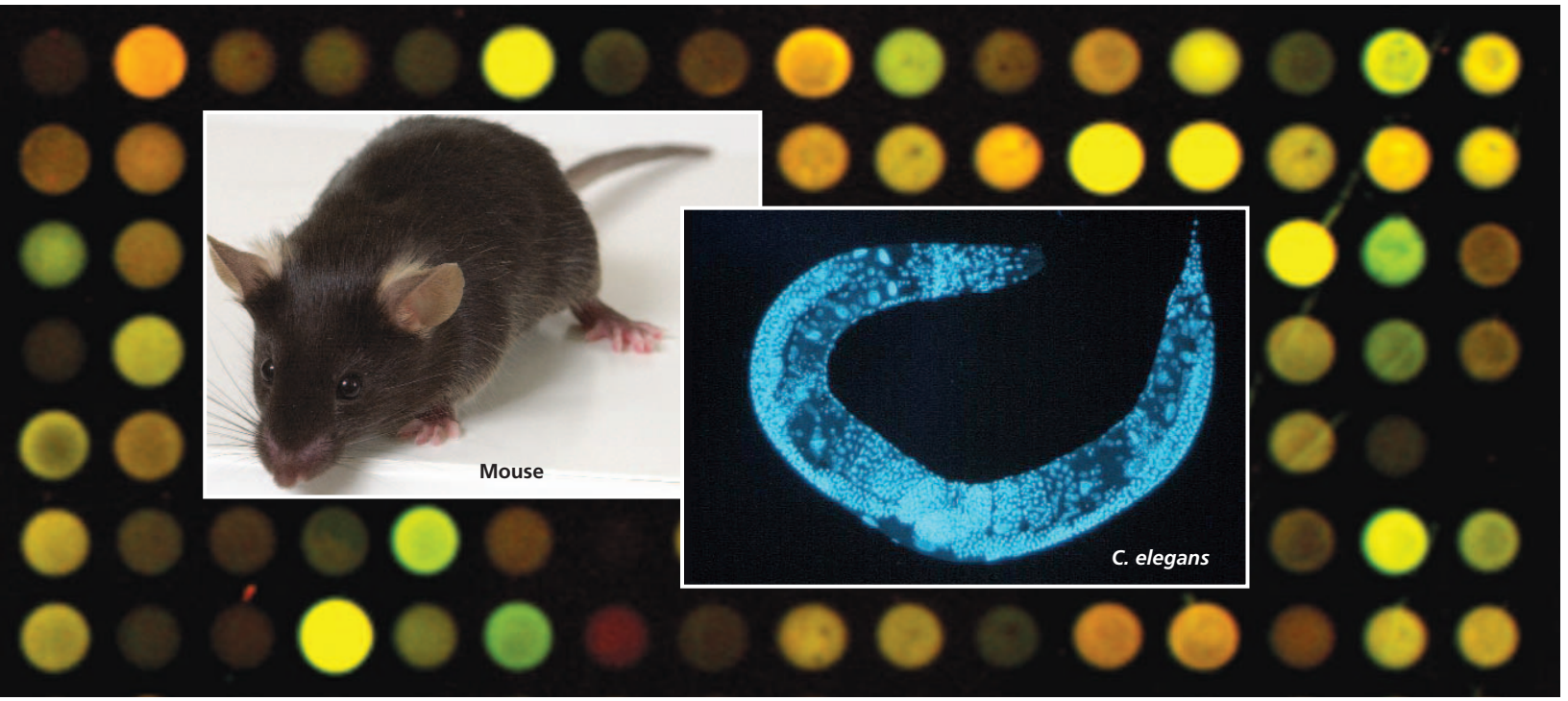
Jonathan Freedman, who heads up the center's Toxicology Core, is the group's *C. elegans* specialist. He is currently developing a high-throughput worm genome–wide microarray to screen compounds and develop signature profiles of response to a variety of toxicants, including bacterial toxins, chemicals, metals, and alkylating agents. He believes that such genomewide screening is the best way to extract useful information. For example, he says, his group has identified about 300 genes that are upregulated in response to cadmium. "But we're not going to go through and look at each gene," he says. Instead, he plans to look at the whole genome to study why cadmium affects all 300 genes, then link that information to what cadmium can do to cause cancer and other diseases.

Evolution of a Field

Freedman is enthusiastic about the Duke center's comparative genomics focus. "We definitely think that's the way toxicogenomics needs to evolve," he says. "There's just so much power in a lot of these alternative species, especially yeast and *C. elegans* and zebrafish. You can do rapid genetics, very rapid RNAi types of studies, and you can do a lot of linking to the genome."

Besides his own work with *C. elegans*, Freedman is also working collaboratively with other member institutes on microarray development and standardization, a crucial issue within the field. Freedman and colleagues distributed material to each of the labs to be used in their microarrays, to see whether all the experiments would come out with the same results. Freedman reports that the first phase of this cooperative work has been completed and is being prepared for publication. The study addressed the issue of data reproducibility by standardizing gene expression experiments across different labs and microarray platforms. The TRC will continue its collaborative efforts at standardization with a new project, possibly addressing comparative genomics.

There should be very exciting results emerging from the Duke center's work in the near future, and from the entire TRC's efforts as well. Schwartz cites three reasons that the TRC initiative will ultimately prove to be of great importance to toxicogenomics: "One is that these approaches will undoubtedly allow us to identify early responses to toxic agents in the environment, and potentially identify individuals before they develop disease. Secondly, using genomics as a way of targeting genes may allow us to short-circuit and hasten the process of identifying which genetic variants are related to susceptibility to environmental agents. And thirdly, this effort will allow us to more clearly phenotype diseases into biological categories of disease, as opposed to clinical or physiological categories of disease that oftentimes lack precision." **–Ernie Hood**



Insets, l–r: Arnold Greenwell/EHP; National Human Genome Research Institute