Journal of Environmental Sciences Vol.5, No.3, pp.253－268,1993 ISSN1001－0742

Prospects for the continued development of environmentally-realistic toxicity tests using microorganisms

John Cairns, Jr.

Department of Biology and University Center for Environmental and Hazardous Materials Studies, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0415, U.S.A.

Paul V. McCormick

Department of Biology, Elizabethtown College, Elizabethtown,Pennsylvania-17022-2298, U.S.A. and Senior Research Associate, University Center for Environmental and Hazardous Materials Studies, Virginia Polytechnic Institute and State University, Blacksburg, Virginia-24061-0415, U.S.A.

Scott E. Belanger

Stream Ecosystem Research Section, Environmental Safety Department, The Procter and Gamble Company, Cincinnati, Ohio-45217, U.S.A.

Abstract — The focus of environmental regulations has changed significantly since the introduction of the bioassay as a standard means of assessing environmental impact. Prominent in this change is an increasing emphasis on protecting the integrity of natural ecosystems, which incorporate community- and system-level properties as well as organismal and population processes. Consequently, support for the use of multispecies testing has widened to include not only ecologists in academia but environmental scientists in the regulatory and industrial sector as well. The reason for this trend is clear: the additional environmental realism gained from tests utilizing communities of organisms allows for greater insight into the potential hazard of chemicals and other forms of human activity to natural ecosystems that cannot be obtained from single species tests alone. Many of the problems cited for multispecies testing early in their evolution as a hazard assessment tool have been refuted or overcome. In particular, the use of natural microbial communities minimizes several shortcomings typically associated with multispecies toxicity testing. This discussion includes the utility of microcosm and mesocosm tests using aquatic microbial communities as hazard assessment tools in conjunction with accumulating information on their performance in toxicity testing protocols. An increasing body of experimental evidence supports an expansion in the use of these tests for a

variety of regulatory and research purposes. A shift in research focus is needed, however, to answer remaining questions and further refine standard protocols for these valuable ecotoxicological tools.
Keywords: environmentally-realistic toxicity tests; microorganisms.

1 Introduction

Concern among scientists over the manner in which the effects of chemicals on natural ecosystems are assessed has been persistent and increasing over the past decade (National Research Council, 1981; O'Neill, 1981; Koeman, 1982; Cairns, 1983; Kimball, 1985; Cairns, 1987; Cairns, 1989; Levin, 1989; Sheehan, 1991). In particular, the extent to which simplistic laboratory tests using a few "surrogate" species are capable of providing accurate predictions of the effects of chemicals on natural ecosystems has been questioned. This debate arose, in part, from efforts to develop a comprehensive strategy for estimating hazard. This strategy, in which one of us (Cairns) was intimately involved, had as its focus the development of a capability for predicting expected environmental concentration of a chemical through an analysis of its expected fate once introduced into an ecosystem (i.e., partitioning, transformation and persistence) and linking such predictions to the expected concentration producing no-adverse-biological effects in the ecosystem (Fig.1).

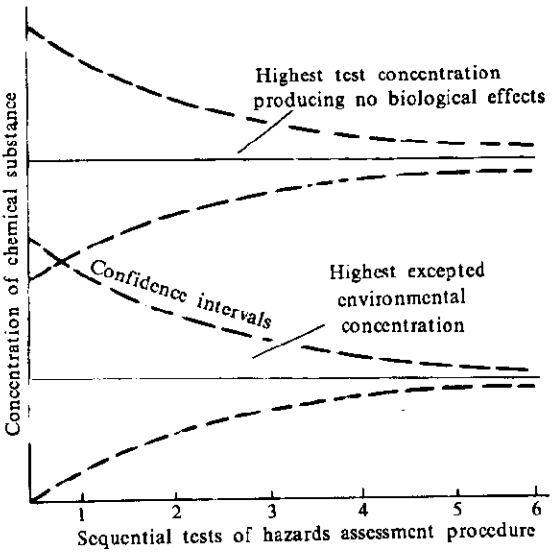

Fig. 1 Sequential tests of hazard assessment procedure

Predictions of either concentration (environmental exposure or biological effect) inevitably involve a certain degree of uncertainty (denoted by confidence intervals in Fig. 1). The level of uncertainty is a function not only of the number of tests conducted, as diagrammed in Fig. 1, but the type of tests conducted as well. In addition to providing precise and reproducible results, it is crucial that laboratory bioassays are also accurate, that is, capable of producing information that can be used to predict possible consequences in the real world.

There are several reasons to expect, on purely ecological grounds, that the response of natural ecosystems to stress is not reliably predictable from responses elicited by test species isolated in laboratory environments. The operation of natural ecosystems is dependent upon process occurring at several levels of biological organization, most prominently those of the individual, population, community and system. Processes at higher levels, such as that of the community (e.g., competition, predation and other trophics interactions) and the ecosystem (e.g., nutrient spiralling), integrate the dynamics of many individual populations and, as such, cannot be modeled merely by understanding toxic effects on a few constituent species. Standard laboratory bioassays typically measure responses of "surrogate" populations to stress but do not directly assess effects on community and ecosystem dynamics. Empirical data are currently lacking to determine the reliability with which laboratory, population-level responses to chemical stress can be extrapolated to predict responses of communities and ecosystems.

2 Resistance to the incorporation of multispecies tests into hazard assessment procedures

Despite the shortcomings of single species tests just discussed, the adoption of more environmentally realistic tests using community- and ecosystem-level end points by regulatory agencies and industry has been relatively slow for a variety of reasons. Despite a strong scientific basis for the use of such tests in hazard assessment protocols (Cairns, 1989), several concerns related to standardization and interpretation have been voiced (Mount, 1985; Tebo, 1985). In short, transformation of the multispecies bioassay from a research tool to a standardized regulatory tool is contingent upon the resolution of two overlapping issues: (1) methodological questions regarding end point selection and measurement and their statistical reliability (i.e., replication and reproducibility); (2) the degree to which such tests increase the environmental realism of hazard assessment protocols and the extent to which such increases translate into increased accuracy in the hazard assessment process.

Criticism of multispecies tests has centered on four specific contentions: (1) there are no generally agree upon community- and ecosystem-level end points for use in multispecies testing; (2) end points that are used in multispecies tests are not decisive; (3) predictions based on multispecies tests are no more accurate or precise than those based on single species tests; (4) multispecies tests are not cost-effective.

3 The utility of multispecies tests in practice: evidence from tests with microbes

The microbial community is a ubiquitous feature of aquatic ecosystems and provides a useful tool for directly assessing the hazard that chemicals pose to natural ecosystems. Several taxonomic groups comprise the microbial community, including protozoa, algae, bacteria, several groups of microinvertebrates, and certain fungi. Species interactions within and between these groups are varied and rival in complexity those occurring in communities of higher organisms (Cairns, 1977; McCormick, 1991). The range of sensitivity of microbes to stress is comparable to that of higher organisms as well (Patrick, 1968).

There are several distinct advantages to the use of microbial communities for hazard assessment. Unlike most higher organisms, many microbe species appear to exhibit a global distribution and consequently, can be found in appropriate habitats around the world. The microbial community comprises the bulk of the biomass in many aquatic ecosystems and fulfill crucial roles in energy flow pathways and nutrient regeneration. The typical microbial community is composed of hundreds of species that vary greatly in their individual sensitivity to chemical stressors, an attribute that ensures that the underlying tolerance distribution of randomly selected assemblages should be similar and that the extreme sensitivity and tolerance of a few species should not markedly skew test results. Protocols using microbes can usually be conducted more rapidly and in smaller test containers than for protocols using higher organisms, thus making them a relatively cost-effective means of evaluating community- and ecosystem-level responses to anthropogenic stressors.

Two avenues of research have been pursued in developing multispecies tests using microbial communities: (1) testing with gnotobiotic, or species-defined, assemblages constructed on the basis of ease of culture in the laboratory and representativeness in the aquatic food web; (2) collection of microbial communities indigenous to natural ecosystems for use in laboratory tests. Gnotobiotic test communities offer much more control for the investigators and, thus, are more amenable to standardization than test designs utilizing indigenous communities. However, while gnotobiotic tests strive to recreate important attributes of natural ecosystems, this assumption is rarely verified. Thus, as with single species tests, several questions relating to the correspondence between results in the laboratory and natural test systems are implicit in the design of gnotobiotic test systems (Giesy, 1985). Use of indigenous microbial communities greatly increases the likelihood that important community- and

ecosystem-level processes will be operating in the test system. While extensive testing has not been pursued to determine the degree of reproducibility of such tests between trials and ecosystems, various lines of evidence suggest that reproducibility in community organization and response to stress are reasonably good (Pratt, 1985a; 1989; Niederlehner, 1990).

Consequently, the remainder of this discussion is devoted to a synthesis of the mounting evidence to support the use of toxicity tests with indigenous microbial communities in hazard assessment protocols.

Current research in laboratories at Virginia Tech and Procter and Gamble has investigated the process of microbial community development on artificial substrate as a means of evaluating the toxicity of different chemical agents. Use of artificial substrates minimizes several problems associated with the collection of indigenous microbial communities on natural surface, including variability in biomass and species composition among natural collections as result of microsite differences in the physico - chemical environment (Cairns, 1982). Polyurethane foam blocks immersed in an aquatic ecosystem provide a suitable habitat for the development of the indigenous microbial community (Cairns, 1982; Stewart, 1985; McCormick, 1988). Communities that closely resemble those found on the natural substrate typically develop on artificial substrates within a few weeks following their placement in an aquatic ecosystem.

Polyurethane foam units (PFUs) and their associated microbial community are easily transported to the laboratory for various testing purposes. Effects of exposure to a chemical agent may be evaluated directly on the PFU community transported into the laboratory or by measuring effects on the process of microbial community development on initially unpopulated or "island" PFUs in laboratory test containers in the presence of these source or "epicenter" substrate. This colonization process adheres to the MacArthur-Wilson equilibrium model (MacArthur, 1967) with regard to patterns of species accrual:

$$St = S_{eq}(1-e)^{-Gt} \quad , \tag{1}$$

where St is the number of species present on the substrate at time t, S_{eq} is the number of species at equilibrium, and G is the rate of species accrual. The details, consequences and limitations of this particular model of microbial colonization are discussed elsewhere (Cairns, 1979; 1982; Pratt, 1985b; McCormick, 1988; 1990) and will not be repeated here.

Effects of stress on the process of community development have most frequently been assessed by measuring dose-related changes in the rate (G) and extent (S_{eq}) of

species accrual on "island" substrates or the number of species maintained on "epicenter" substrates. While these measures have been extremely useful in characterizing toxicity, gathering necessary data requires time and expertise since these measures require taxonomic identifications. A varicty of other structural (e.g., total biomass measures) and functional (e.g., enzymatic activity) responses, most of which require less time and expertise to measure, have been evaluated on occasion (Table 1).

Table 1 Commonly used parameters of microbial community structure and function suitable for measurement in laboratory microcosms

Structure
Species richness
Community composition (e.g., similarity indices)
Community biomass (e.g., chlorophylL−a, AFDM, total protein, etc.)
Colonization rate
Function
Heterotrophic index
Trophic complexity
Primary productivity (e.g., oxygen evolution, ^{14}C-uptake)
Community respiration (e.g., oxygen uptake, electron transport activity)
Productivity:respiration
Productivity:biomass/Respiration:biomass
Enzymatic (e.g., alkaline phosphatase) activity
Substrate processing

This test design has evolved to the extent that a standard protocol has been proposed that covers preparation, incubation, and measurement (Pratt, 1990).

Hazard assessment protocols encompass as many as four successive tiers of biological testing (Kimerle, 1978): (1) screening or range finding tests used to determine acute toxicity to a small array of surrogate species and provide sufficient information for a preliminary safety evaluation; (2) predictive tests designed to estimate toxicity with maximum accuracy and provide a maximum allowable toxic concentration (MATC) for environmental exposure; (3) confirmative tests to evaluate environmental effects under anticipated conditions; and (4) monitoring following introduction of the chemical into the environment to provide an ongoing error control mechanism. The objective at each successive stage is to further increase confidence in the level of toxicity. It follows, therefore, that the complexity, duration, and environmental realism of tests increase as testing proceeds, along with the associated cost. Variations of the microbial test design just described have contributed to an understanding of environmental

hazard at each of these tiers of testing.

3. 1 Screening tests

Microbial community tests of the general type just described have most frequently been used as part of predictive or confirmative tiers of the hazard assessment process, and concentrations used in such tests are usually selected based on single species screening tests. In a few instances (Niederlehner, 1985; McCormick, 1986), short-term (e.g., 48h) rangefinding experiments have been conducted using the dose-response relationship of epicenter PFUs to stress to predict the degree of toxicity. Microbial communities on pre-colonized PFUs have consistently been extremely resistant to the effects of chemical stressors compared with most convential single species screening tests and parameters (e.g., rate and extent of species accrual) measured in predictive and confirmative tests (Niederlehner, 1985; McCormick, 1986), thus limiting their utility as a rangefinding tool.

Measurement of effects of chemical stress on the initial phases of colonization (e.g., the number of species accrued on initially unpopulated PFUs over a specified period of time) provides a more suitable screening tool than the dosing experiments just described. Pratt and colleagues (Pratt, 1988a) used a 7- day laboratory colonization test to evaluate the toxicity of eluates leached through soils collected from several hazardous waste sites. Sites were ranked for their relative toxicity using the EC20 of this test (i.e., the concentration of eluate that elicited a 20% reduction in the number of colonizing species compared with an undosed control) and the EC50 of a conventional acute single species test using the cladoceran Daphnia *magna*. Rankings were generally similar between the two tests (Table 2).

Table 2 The toxicity of putative hazardous waste sites ranked using both single species and multispecies test procedures
(See text details, reprinted from Pratt, 1988a, with permission from Kluwer Academic Publishers)

Type of waste	Single species		Multispecies		Rank sum
	EC50	Rank	EC20	Rank	
Wood preservatives	<0.5%	1	2.2%	2	3
Wood preservatives	<1.0	2.5	0.035	1	3.5
Heavy metals	<1.0	2.5	2.7	3	5.5
Coal leachates	50.1	4	—[a]	4	8
Heavy metals	79.4	5	NT[b]	6	11
Solvents	NT	6.5	NT	6	12.5
Control	NT	6.5	NT	6	12.5

a. Toxicity shown but dose-response relationship not significant; b. not toxic

Although modest differences were observed in some instances, possibly due to degradation of toxic material or adaptive mechanisms operating in the community-level tests. While results such as these indicate the need for multiple test procedures for precise rangefinding, they do not suggest that evidence from multispecies tests is absolutely necessary for the purposes of bracketing an effect concentration, especially when the taxonomic expertise required to conduct such tests is considered.

3. 2 Predictive microcosm tests

Microcosm tests with indigenous microbial communities have been used extensively as predictive tools for toxicity assessment (Table 3). Ranges of valus for multispecies tests are based on responses of varying numbers of structural and functional parameters in the same test.

Table 3 Summary of findings of predictive toxicity tests using the microcosm design described in this article, and correspondence between these findings and those relying on single-species tests

Toxicant	Estimated toxicity, μg/L	Comparison with standard lab tests	Field validation
Cadmium	MATC=0.20	Acute MATC=42 Chronic MATC=0.82	NA
TFM	MATC<100μg/L (stimulatory)	LC25 for several fish Species: 5000−44000	NA
Copper	LOEC<6.6−36.5 MATC<6.6−26.7	LOEC for several species: 6.1−60.4, chronic MATC=8.2	NA
Zinc	LOEC<4.2−89.2 MATC<4.2−51.6	Average LOEC for several Species=47, chronic MATC=47	NA
Chlorine	LOEC<2.1−261 MATC<2.1−176	Average LOEC for several Species=3.4, chronic MATC=11	NA
Atrazine	MATC=17.9−193	MATC=71−3,400	Field mesocosms MATC=6−144
Effluent	LOEC=1%	Acute daphnids and fish LC50: 18.8−63.1% Chronic daphnids:1% (using an 1% application factor)	Stream inverte−brate responses predicted by lab tests
phenol	LOEC=300−23,300+	Several chronic single Species LOEC=2600	NA
Ammonia	LOEC=10−430 MATC<10−260	Several chronic single species Species LOEC=2−612,MATC=28	NA
Selenium	MATC=14.4	Daphnid and fish MATCs= 180−360	Lotic field mesocosm MATC=17.7

'LOEC=lowest observable effect concentration; MATC=maximum allowable toxic concentration (Niederlehner, 1985, 1986; McComick, 1986; Pratt, 1987a; Pratt, 1987b; Pratt, 1988c; Pontasch, 1989; Pratt, 1989; Pratt, 1990)

Although not a review of details of the results of each of the cited studies, Table 3 does, however, provide a summary that indicates several important points. First, a comparison of the results of the multispecies test used in these studies with those generated using standard laboratory protocols indicates that no consistent correspondence between responses at different levels of biological organization can be assumed. Multispecies tests are not consistently more sensitive than single species tests as some advocates of multispecies tests once argued. However, effects on community- and ecosystem-level processes are frequently detected at chemical concentrations that have been deedmed "safe" (i.e., associated with an acceptable risk) by the USEPA largely on the basis of effects data from single species tests. This result indicates a need to incorporate tests at different levels of biological organization into the predictive tier of testing.

Our contention that laboratory toxicity data cannot be reliably extrapolated among levels of biological organization differs from that of other researchers (e.g., Sloof, 1986). Previous assessments have encompassed population and ecosystem responses to several toxicants and, while they indicate that, in general, different levels of biological organization are similar in their sensitivity to stress, the validity of developing a generic predictive relationship of toxicity among levels is by no means convincing. Statistically significant relationships between single species and multispecies data across toxicants, such as those derived by Sloof and colleagues (Sloof, 1986), include several cases where predicted effect concentration derived at different levels of biological organization are similar, and fewer (i.e., about 15% of the total as derived from the database of Sloof, 1986) instances where discrepancies among levels of biological organization are considerable (i.e., greater than an order of magnitude). While the presence of a few "outliers" may not unduly affect the overall significance of a statistical relationship between different types of tests, the most compelling question from an environmental standpoint should be whether current predictive capabilities, which may be in gross error as much as 15% of the time, are acceptable.

A second, equally important conclusion of studies presented here (as well as others summarized in Cairns and Niederlehner, 1987) is that multispecies tests provide information on the nature of environmental impact that would rarely be indicated by standard single species test results. In general terms, this contention is illustrated by experimental data documenting the variability in the relative sensitivity of ecosystem structure and function to different forms of chemical stressors. For example, while structural changes were judged more sensitive to certain types of effluents (Shen, 1986), functional attributes may be much more sensitive to other toxicants, such as the herbicide atrazine (Pratt, 1988b). Given that there is no consistent dose-response

relationship among parameters within a given level of biological organization, how can such a relationship be assumed between levels of biological organization in order to protect natural ecosystems? In general, then, claims of "generic" extrapolation methods must be questioned from a scientific viewpoint.

An essential requirement for toxicity tests at any tier of the hazard evaluation process is that the results be reproducible among trials. Such evidence is accumulating both for responses determined using communities collected from the same ecosystem at different time (Niederlehner, 1990) and from different ecosystems at similar times (Pratt, 1989). More information in this area would certainly be useful.

3. 3 Confirmative studies

The ultimate interest in any laboratory test sequence is the degree to which results accurately predict impact in real ecosystems. Such evidence is essential for determining the utility of various tests, but, unfortunately, laboratory estimates of toxicity are rarely subjected to field validation. Results of some of the predictive laboratory tests described above have been subject to field validation using experimental and observational studies (Table 3). These sutdies suggest that the relative predictive capability of multispecies tests may be sometimes greater than that of standard single species tests.

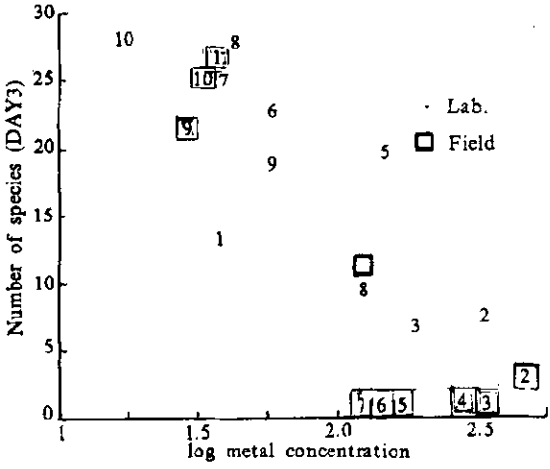

Fig. 2 Relationship between the numbers of species and metal concentration

The use of microbial colonization protocols for providing realistic estimates of environmental hazard is illustrated by a study of Shen and co-workers (Shen, 1986), who performed a series of predictive and confirmative tests using protozoan communities to assess the effects of complex effluents on stream ecosystems. Results of this work were some of the first to judge the accuracy of microcosm tests using these communities in predicting environmental effects. Laboratory microcosm tests measuring protozoan colonization predicted an IC20 for species accrual(i.e., a 20% reduction in species number of 48 μg/L for metals contained in the effluents), a prediction that corresponded closely to that found in field trials (51 μg/L, Fig.2). Structural (i.e., species richness) attributes were especially useful in trials predicting deleterious effects in the natural ecosystem.

A recent undertaking by the Procter and Gamble Company (Cincinnati, Ohio, USA) provides an example of confirmative studies using microbial communities by industry. Procter and Gamble is one of the largest manufactures of household consumer products in the world. As such, many of their products are disposed "down the rain" and enter natural receiving systems via municipal treatment facilities. In order to evaluate potential effects of chemicals to stream communities under environmentally realistic conditions, an experimental stream facility (ESF) was constructed to augment standardized laboratory and field testing protocols (Woltering, 1989).

The ESF consists of replicate indoor stream channels, each consisting of a series of flow environments including a 4.3m run (2.5cm deep with a current velocity of 45－60 cm/s), and a pool environment (90 cm deep with a holding volume of 568 L; Belanger, 1992). Channels are supplied with a 9:1 mixture of water from an adjacent stream, rated as an exceptional warmwater fishery, and final effluent taken immediately prior to chlorine disinfection from a sewage treatment facility just downstream of the ESF to provide an environment similar to that in a natural stream receiving a typical level (~ 10% by volume) of treated effluent containing household chemicals. Water is delivered at a rate and force that allows natural microbial and macroinvertebrate colonization to oc ur and is not recirculated in order to simulate ambient stream conditions as closely as possible. Metal halide lamps provide full sunlight spectrum lighting at approximately 68 μmol/(s· m^2) photosynthetically active radiation (PAR). While recreating several aspects of environmental realism, the design of these stream channels also allows for certain elements of control necessary for precisely determining toxic effects (e.g., maintenance of a constant current velocity and continuous dosing of test materials to maintain consistent levels of environmental exposure).

The performance of the streams at the ESF was recently evaluated by conducting a toxicity test using a model compound, the surfactant C12-TMAC (dodecyl trimethyl ammonium chloride), the environmental hazard of which has been extensively evaluated (Woltering, 1989). Four channels completed at the time were used for experimentation: two channels received no C12-TMAC (controls) while two other channels were dosed continuously with a modest (50 μg/L) and relatively high (250 μg/L) dose of C12-TMAC, respectively. PFUs were placed in the pool of each channel either 28 days prior to dosing to develop a mature stream microbial community in order to evaluate community resistance to stress or on day 0 of dosing to evaluate the effects of the surfactant on the process of microbial community development.

Community structure, including the taxon richness and composition of both immature and mature protozoan and diatom assemblages, was generally unaffected by even

high concentrations of the surfactant. Modest changes in other parameters, including diatom growth processes and protozoan trophic relationships (i.e., the number of taxa in different feeding groups; Pratt, 1985b) were indicated within the range of concentrations tested, but, were not consistent (progressive and/or dose-dependent) throughout the experiment.

Indigenous microbial communities in the ESF channels appeared to be relatively insensitive to the surfactant compared with surrogate species tested in standard laboratory experiments. A lack of statistical power probably contributed to the rather poor sensitivity of these PFU communities to the chemical. However, realistic variation in environmental factors may have further contributed to discrepancies between estimates of hazard generated from standard laboratory tests and effects observed in the stream mesocosms. This view is supported since other biological components of the streams, such as periphytic algae on hard substrate and invertebrate responses, were entirely consistent with field-based investigations of TMAC toxicity (Woltering, 1989; Lewis, 1986).

3.4 Monitoring

Undoubtedly, the most neglected aspect of environmental hazard assessment is the implementation of a monitoring program to track the fate and effects, as well as exposure, of a compound once it has been released into the environment. Monitoring programs should account for cumulative effects, environmental variability, and other factors that are rarely considered in laboratory protocols. As such, monitoring functions principally as an ongoing error detection and correction loop, which validates laboratory predictions and, when unexpected impacts are detected, can be used to recommend corrective actions. Cost is certainly a major concern at this stage of testing, given the possible temporal and spatial scope of monitoring efforts.

3.5 Future prospects and needs

We have discussed the need for continued development of environmentally realistic toxicity tests, an specifically, the usefulness of microbial community tests for this purpose. Microbial community test systems provide a scientifically valid means of assessing potential toxicant impacts on community- and ecosystem-level attributes of aquatic systems. There is increasing evidence that such test exhibit an adequate degree of replication and reproducibility for routine predictive purposes. Futhermore, methodologies are becoming standardized (Pratt, 1990). These features, combined with savings in time and money compared with many other multispecies and standard single species tests, make prospecs for increased use of microbial microcosm tests quite good.

Based on the current status of development, we can evaluate three alternative roles for multispecies tests, particularly those utilize microbial communities:

3.5. 1 Limited to use as an exploratory research tool

Experimental ecosystems have been invaluable tools in advancing an understanding of many fundamental ecological processes. However, is the usefulness of these test systems in toxicological research limited to the satisfaction of intellectual curiosity? This was the contention of many detractors years ago when these tests were first proposed as a means of predicting toxicity. Even at that time, a reasonable amount of evidence supported the development of microcosm community- and ecosystem-level tests as regulatory and exploratory tools (Giesy, 1980). Evidence presented in this discussion, as well as evidence from other researchers (Cairns, 1986), provides a substantial amount of additional support for a dual role for community- and ecosystem-level tests, in particular those using natural microbial communities. These tests have been effective not only as means of understanding the mechanisms that drive community and ecosystem responses to stress, but also for the routine assessment of hazard as well.

3.5.2 Expand used to include optional multispecies tests at higher tiers of testing

This discussion has shown that multispecies tests provide information that can be reliably used to : (1) validate environmental response thresholds predicted from less realistic single species tests, and (2) provide information regarding the nature of potential environmental impacts that cannot possibly be obtained consistently from single species tests. Such information will certainly increase the accuracy of predictions regarding environmental hazard in many instances. But, in what capacity can they be most effectively used? Variations of the microcosms tests we describe appear to show promise as screening tests (Ranking the toxicity of different effluents or other hazardous wastes). Most evidence to date supports the use of these tests at predictive and confirmative tiers of testing. We argue that microcosm tests such as those described here have already proven themselves useful as a hazard assessment tool at these levels, and further integration into standard protocols is thus warranted.

3.5.3 Expand to routine use at all tiers

It is not completely clear at this time the extent to which multispecies tests will be needed to augment and/or replace standard single species testing in hazard assessment schemes. Undoubtedly, single species testing will remain the backbone of hazard assessment protocols for the foreseeable future. Methodolgical and statistical considerations are well understood for most standard single species tests. A similar sort of systematic database will be required for promising multispecies tests, along with further evidence for their use in decision-making, before radical shifts can be expected in the focus of current regulatory requirements and testing in laboratories where equipment and staffing center largely on single species. We expect that the impetus

for such change must come largely from the academic and industrial sectors rather than from regulatory agencies.

We have moved beyond the point where credible evidence for the use of multispecies test results is needed and have gained substantial insight into many problems related to replication and standardization. There remains a need for a broad database with which to compare the predictive capabilities of multispecies and single species tests. Obviously, this will require both predictive and confirmative studies as well as monitoring to test for the efficacy of long- term predictions based on laboratory tests. From a methodological standpoint, perhaps the most urgent need at this time is evidence for reproducibility, both among tests using microbial (or other taxonomic) communities from the same ecosystem at different times, to assess the extent of temporal (seasonal) variability, and among ecosystems to gauge the extent to which microbial communities vary in their sensitivity to anthropogenic stress. This information will be extremely useful not only for standardizing laboratory protocols but for determining the extent of variability that should be expected in nature. This latter problem has rarely been addressed in any detail and can most effectively be answered using multispecies tests with a cosmopolitan group of organisms, such as microbes.

4 Conclusion

At this juncture, it is clear both that measurable progress has been made in the development of environmentally realistic hazard assessment procotols that include community- and ecosystem-level end points and that a great deal of "standardization" work is still required before anything approaching an immutable case can be made for their routine incorporation into hazard assessment procotols. It is somewhat surprising that more information from a broader cross-section of the scientific community has not been accumulated, given that many of the basic reservations concerning the incorporation of multispecies tests into regulatory protocols were originally voiced a decade or more ago. Future work in this area will need to address problems such as that of reproducibility in a systematic way to develop multispecies tests as effective hazard assessment tools. These efforts will undoubtedly be most fruitful if they include collaborative efforts between academia, industry and regulators.

Acknowledgements — This manuscript and specific experiments discussed within were sponsored, in part, by grants from The Procter & Gamble Company and the du Pont Educational Foundation, although the authors take full responsibility for the views expressed herein. The authors thank Darla Donald for editorial assistance in the

preparation of this manuscript for publication and Teresa Moody for putting the dictation on the word processor.

References

Belanger SE, Barnum JB, Woltering DM, Bowling JW, Ventullo RM, Schermerhorn SD, Lowe RL. Algal periphyton structure and function in response to consumer chemicals in experimental stream mesocosms, Proceedings of the aquatic simulated fiels studies symposium (Ed. By RL Graney, JH Rodgers, J Kennedy). Chelsea, Michigan: Lewis Publishers (In press)

Cairns JJr. Aquatic microbial communities. New York: Garland Press, 1977

Cairns JJr. Artificial substrates. Ann Arbor, Michigan: Ann Arbor Science Publishers, 1982

Cairns JJr. Hydrobiologia, 1983; 100:47

Cairns JJr. Community toxicity testing, American Society for Testing and Materials, Philadelphia, Phennsylvania, 1986

Cairns JJr, Henebry MS. Artificial substrates (Ed. by J Cairns Jr). Ann Arbon, Michigan: Ann Arbon Science Publishers, 1982: 23

Cairns JJr, Niederlehner BR. Hydrobiologia, 1987; 153: 87

Cairns JJr, Pratt JR, Hydrobiologia, 1989; 188/189:5

Cairns JJr, Dickson KL, Maki AW. Estimating the hazard of chemical substances to aquatic life, STP 657 (Ed. by J Cairns Jr, KL Dickson, AW Maki). Philadelphia: American Society for Testing and Materials, 1978: 191

Cairns JJr, Kuhn DL, Plafkin JL. Methods and measurements of attached microcommunities: a review (Ed. by RL Weitzel). Philadelphia: American Society for Testing and Materials, 1979:39

Giesy JPJr. Microcosms in ecological research, National Technical Information Center, Springfield Virginia, 1980

Giesy JPJr. Aquatic toxicology and hazard assessment: eight symposium (Ed. by RC Bahner, DJ Hansen). Philadelphia: American Society for Testing and Materials, 1985:67

Kimball KD, Levin SA. Bioscience, 1985; 35:165

Kimerle RA, Gledhill WE, Levinskas CJ. Estimating the hazard of chemicals to aquatic life (Ed. by J Cairns Jr, KL Dickson, A.Maki). Philadelphia: American Society for Testing and Materials, 1978:132

Koeman JH. Ecotoxicol Environ Safety, 1982; 6:358

Levin SA, Harwell MA, Kelley JR, Kimball KD. Ecotoxicology: problems and approaches. New York: Springer-Verlag, 1989

Lewis MA, Taylor MJ, Larson RJ. Community toxicity testing (Ed. by J Cairns Jr). Philadelphia: American Society for Testing and Materials, 1986:241

MacArthur RH, Wilson EO. The equilibrium theory of island biogeography, Princeton. New Jersey: Princeton University Press, 1967

McCormick PV, Cairns JJr. Hydrobiologia, 1990; 196:229

McCormick PV, Cairns JJr. Hydrobiologia, 1991; 213:83

McCormick PV, Pratt JR, Cairns JJr. Community toxicity testing (Ed. by J Cairns Jr). Philadelphia: American Society for Testing and Materials, 1986:224

McCormick PV, Pratt JR, Jenkins DG, Cairns JJr. Trands Am Microsc Soc, 1988; 107:259

Mount DI. Multispecies toxicity testing (Ed. by J Cairns Jr). New York: Pergamon Press, 1985: 13

National Research Council. Testing for effects of chemicals on ecosystems. Washington, DC.: National Academy Press, 1981

Niederlehner BR, Cairns JJr. Environ Pollut, 1990; 66:207

Niederlehner BR, JR Pratt, AL Buikema Jr, J Cairns Jr. Environ Toxicol Chem, 1985; 4:155
Niederlehner BR, JR Pratt, AL Buikema Jr, J Cairns Jr. Community toxicity testing (Ed. by J Cairns Jr) Philadelphia: American Society for Testing and Materials, 1986:30
O'Neill RV, Waide JB. Management of toxic substances in our ecosystem (Ed. by BW Cornaby). Ann Arbor, Michigan:Ann Arbor Science Publishers, 1981:43
Patrick R, Cairns JJr, Scheier A. Prog Fish Cult, 1968; 30:137
Pontasch KW, Niederlehner BR, Cairns JJr. Environ Toxicol Chem, 1989; 8:521
Pratt JR, Bowers NJ. Tox Assess, 1990; 5:189
Pratt JR, Cairns JJr. J Protozool, 1985a; 32:409
Pratt JR, Cairns JJr. Hydrobiologia, 1985b; 121:103
Pratt JR, Niederlehner BR, Bowers NJ, Cairns JJr. Tox Assess, 1987a; 2:417
Pratt JR, Niederlehner BR, Bowers NJ, Cairns JJr. International conference on heavy metals in the environment (Ed. by SE Lindberg, TC Huntchinson). Edingburg, United Kingdom: CEP Consultants Ltd, 1987b:324
Pratt JR, McCormick PV, Pontasch KW, Cairns JJr. Water Air Soil Pollut, 1988a; 37:293
Pratt JR, Bowers NJ, Niederlehner BR, Cairns JJr. Arch Environ Contam Toxicol, 1988b; 17:449
Pratt JR, Bowers NJ, Niederlehner BR, Cairns JJr, Environ toxicol Chem, 1988c; 7:679
Pratt JR, Bowers NJ, Niederlehner BR, Cairns JJr. Tox Assess, 1989; 4:161
Scheehan PJ. Fundamentals of the management of toxic chemicals (Ed. by RP Cote, PG Wells). Boston: Urwin Hyman Press, 1991:79
Shen YF, Buikema ALJr, Yongue WHJr, Pratt, JR, CAirns JJr. J Protozool, 1986; 33:146
Slooff W, Van Oers JAM, De Zwart D. Environ Toxicol Chem, 1986; 5:841
Stewart PM, Pratt JR, Cairns JJr, Lowe RL. Trans Am Microsc Soc, 1985; 104:369
Tebo LRJr. Multispecies toxicity testing (Ed. by J Cairns Jr). New York: Pergamon Press, 1985:19
Woltering DM, Bishop WE. The risk assessment of environmental and human health hazards (Ed. by DJ Paustenbach). New York: John Wiley and Sons, 1989:345

(Received April 28, 1992)