SEDIMENT QUALITY AND BIOLOGICAL EFFECTS IN SAN FRANCISCO BAY

BAY PROTECTION AND TOXIC CLEANUP PROGRAM

FINAL TECHNICAL REPORT

California State Water Resources Control Board
Division of Water Quality

San Francisco Bay Regional Water Quality Control Board

California Department of Fish and Game
Marine Pollution Studies Laboratory

California State University
Moss Landing Marine Laboratories

University of California, Santa Cruz
Institute of Marine Sciences

August, 1998

SEDIMENT QUALITY AND BIOLOGICAL EFFECTS IN SAN FRANCISCO BAY

BAY PROTECTION AND TOXIC CLEANUP PROGRAM

FINAL TECHNICAL REPORT

Authors:

John W. Hunt, Brian S. Anderson, Bryn M. Phillips, John Newman, Ron S. Tjeerdema
Institute of Marine Sciences, University of California, Santa Cruz

Karen Taberski
San Francisco Bay Regional Water Quality Control Board

Craig J. Wilson
Division of Water Quality, California State Water Resources Control Board

Mark Stephenson, H. Max Puckett
Marine Pollution Studies Laboratory, California Department of Fish and Game

Russell Fairey and James Oakden
Moss Landing Marine Laboratories

ACKNOWLEDGMENTS

This study was conducted with the assistance of a number of colleagues from the Marine Pollution Studies Laboratory and its facilities at Granite Canyon, Moss Landing, and Santa Cruz, California. The authors would like to specifically acknowledge the following people for their efforts in the successful completion of the studies described in this report.

University of California, Santa Cruz, Staff Conducting Toxicity Testing at Granite Canyon

Steve Clark, Matt Englund, Michelle Hester, Hilary McNulty, Patty Nicely, Steve Osborne, Witold Piekarski, Kelita Smith, Shirley Tudor, Lisa Weetman, Michelle White

University of California, Santa Cruz, Trace Organics Laboratory

Jonathan Becker, Gloria J. Blondina, James Derbin, Linda Hannigan, Deborah Holstad, Dana Longo, Raina Scott, Katharine Semsar, Thomas Shyka, Matthew Stoetling, Else Gladish-Wilson, Laura Zirelli

California Department of Fish and Game, Trace Metal Analytical Laboratory

Jon Goetzl, Gary Ichikawa, Jim Kanihan, Kim Paulson, Mark Pranger

San Jose State University, Sample Collection and Data Management

Ross Clark, James Downing, Michele Jacobi, Eric Johnson, Brenda Konar, Stewart Lamerdin, Eli Landrau, Cassandra Roberts

Moss Landing Marine Laboratories (SJSU), Organic Carbon and Grain Size Analysis

Bill Chevalier, Steve Fitzwater, Craig Hunter, Pat Iampietro, Sean McDermott, Michelle White

Moss Landing Marine Laboratories (SJSU), Benthic Community Analysis

Carrie Bretz, Diane Carney, Jim Oakden, John Oliver, Pete Slattery

We would especially like to thank Michelle White and Patty Nicely for assistance with preparing this report, specifically for their skills with mapping, data compilation, and statistical analysis. We greatly appreciate the editorial review by Gita Kapahi, SWRCB.

This study was funded by the State Water Resources Control Board and San Francisco Bay Regional Water Quality Control Board through the Bay Protection and Toxic Cleanup Program.

EXECUTIVE SUMMARY

Study Objectives

The Bay Protection and Toxic Cleanup Program (BPTCP) was established by the California State Legislature in 1989 with four major goals:

1) To provide protection of present and future beneficial uses of the bay and estuarine waters of California;
2) To identify and characterize toxic hot spots;
3) To plan for toxic hot spot cleanup or other remedial or mitigation actions; and
4) To develop prevention and control strategies for toxic pollutants that will prevent creation of new toxic hot spots or the perpetuation of existing ones within the State's bays and estuaries.

These goals are being addressed through activities in each of the coastal Regional Water Quality Control Boards, including that representing the Bay Region. BPTCP program activities in the San Francisco Bay Region have included initiating the Regional Monitoring Program and conducting studies of fish tissue contamination (as described in the Introduction section of this report). The BPTCP has also implemented regional monitoring studies to identify toxic hot spots. The four major objectives of BPTCP monitoring in the San Francisco Bay Region, as described in this report, are:

1) To identify locations in enclosed bays, estuaries, or the ocean that are potential or candidate toxic hot spots;
2) To determine the extent of biological impacts in portions of enclosed bays and estuaries not previously sampled (areas of unknown condition);
3) To confirm the extent of biological impacts in enclosed bays and estuaries that have been previously sampled; and
4) To assess the relationship between toxic pollutants and biological effects.

The focus of BPTCP monitoring in San Francisco Bay has been to conduct sediment quality assessments in several phases: 1) Approximately 100 reports were evaluated for previous information on water and sediment quality; 2) A large number of bay and wetland sites were surveyed in the Pilot Regional Monitoring Program (PRMP), which also included a methods validation study along a pollution gradient; 3) A reference site study was completed that evaluated ambient conditions in the Bay, and evaluated toxicological and statistical methods for differentiating polluted sites and reference conditions, 4) Approximately 127 stations from throughout the region (selected on the basis of previous information and PRMP results) were screened for sediment toxicity and/or chemistry; and 5) A number of sites that exhibited toxicity and/or elevated chemistry were resampled for additional biological and chemical analyses to confirm previous results. This confirmation survey incorporated three components commonly known as the sediment quality triad: toxicity testing, chemical measurement, and benthic community analysis. Additional samples were collected at selected confirmation sites to estimate the bioavailability of sediment-associated chemicals. Concurrent with this phased sediment monitoring effort, a study was conducted in 1994 to determine chemical concentrations in fish tissues. The results of that study were the basis for a subsequent public health advisory for fish consumption in the Bay and Delta.

Tasks Accomplished

This report describes the results of BPTCP sediment monitoring activities in the San Francisco Bay Region to identify toxic hot spots. During the screening phase of this study, 127 sites that had been identified in previous investigations were screened for sediment toxicity. Since funding constraints precluded comprehensive assessments at each screening site, toxicity testing was used as the screening tool. Toxicity tests are direct, precise indicators of the integrated effects of sediment contaminants, and they provide information about biological impacts of pollutants, information difficult to discern solely from chemical measurements. Generally, two toxicity tests were used at each screening site: a solid-phase sediment test with benthic amphipods, and a sediment porewater test using developing embryos of sea urchins. As methodological improvements were incorporated during the study, some screening samples were tested with sea urchins exposed to the sediment-water interface instead of porewater.

After reviewing the screening data and information from previous studies, a number of stations were resampled during the confirmation phase of the study. Twelve stations were resampled and analyzed with the sediment quality triad, including two toxicity tests, sediment chemistry, and benthic community analysis. Ten of these stations were also analyzed for bioaccumulation, using 28-day laboratory exposures with the clam *Macoma nasuta*. A total of 46 stations were screened for a broad suite of trace metal and organic compounds, and a total of 143 samples were analyzed for mercury and PCBs, chemicals that were identified as elevated in fish tissues in the Bay (Fairey et al., 1997) and were the subject of a fish consumption health advisory. An additional 15 stations were resampled and tested with sea urchin larvae in sediment-water interface exposures, because their screening samples exhibited toxicity only in sea urchin porewater tests that were accompanied by elevated sulfide or ammonia concentrations.

In order to provide additional information about potential toxic hot spot sites, linear transects (gradients) were sampled at some confirmation stations to evaluate relationships between sediment chemistry and biological effects at these sites. Phase I sediment Toxicity Identification Evaluations (TIEs) were conducted at two sites, and an abbreviated sediment-water interface TIE was conducted at a third site to investigate possible causes of sediment toxicity.

Major Findings

After screening 127 stations from throughout the Bay area, and returning to 12 of those for more intensive analysis during the confirmation stage, this study successfully identified several highly polluted locations that exhibited adverse biological effects. The study also indicated that 21% of all samples tested were toxic to amphipods, 31% of porewater samples were toxic to sea urchin embryos, and 33% were toxic to sea urchin embryos exposed at the sediment-water interface. Statistical analyses indicated a number of chemicals that were both correlated with biological effects and found at concentrations exceeding sediment quality guideline values.

A number of sites had high concentrations of chemical mixtures, numerous chemicals with concentrations above sediment quality guideline values, and significant biological effects. These sites were categorized based on the magnitudes of chemical concentrations and effects. The sites exhibiting highest chemical concentrations and greatest biological effects included: Stege Marsh, Mission Creek, Islais Creek, Point Portrero (notable for extremely high PCB and mercury concentrations), Pacific Drydock, Castro Cove, Peyton Slough, and San Leandro Bay.

Mercury and total PCBs were identified in a California Office of Environmental Health Hazard Assessment health advisory on consuming fish caught in San Francisco Bay and the Delta. These chemicals were found at elevated concentrations in a number of sediment samples analyzed in this study. PCBs, but not mercury, were accumulated to high levels in clams exposed to 6 of 10 sediment samples tested. Mercury, but not PCBs, was found to correlate with toxicity to sea urchins in sediment-water interface exposures.

In Principal Components Analyses (PCA), sediment quality guideline quotient means (ERMQs) and number of chemicals exceeding guideline values covaried negatively with biological indicators (increasing concentration associated with decreasing biological function). Chemicals identified by PCA that also exceeded guideline values and were significantly correlated with adverse biological effects included: total chlordanes and 2-methylnaphthalene (with amphipod toxicity); cadmium, copper, silver, and zinc (with sea urchin porewater toxicity); and cadmium, copper, and zinc (with sea urchin SWI toxicity).

Sediment quality guidelines, as described in the Methods section, have been derived empirically from a large number of studies nationwide to indicate chemical concentrations often associated with adverse biological effects. The use of guideline values allows simple comparisons of sample concentrations to those observed in numerous other studies. This comparison is useful for perspective, but does not necessarily indicate that chemicals with concentrations above guideline values are responsible for any observed impacts. Only site-specific intensive investigations, using TIEs and other toxicological methods, can be used to determine causal relationships. In the present study, numerous chemicals were found at concentrations exceeding guideline (ERM) values. Of these, chlordanes, PCBs, DDTs, PAHs, dieldrin, copper, mercury, lead and zinc were commonly found above ERMs. Hexachlorobenzene and chlorpyrifos, for which ERM values have not yet been derived, were often found at concentrations above the 90th percentile of the statewide BPTCP sediment chemistry data base. Combined concentrations of chemical mixtures were high at many sites, with 9 sites having mean ERM quotients above the 95th percentile of the statewide distribution.

In tests of 10 samples from the Bay, exposed clams accumulated elevated tissue concentrations of nine chemicals or chemical classes: copper, lead, total chlordanes, total DDTs, dieldrin, total PCBs, LMW PAHs, HMW PAHs, and total PAHs. The identification of these chemicals was dependent on the particular samples tested, the physiology of the

clam *Macoma nasuta*, and the 28-day exposure period of the laboratory tests.

The data provided in this report represent a significant body of information to assist in management efforts to identify and remediate toxic hot spots in San Francisco Bay. A number of sites were identified as having elevated pollutant concentrations and severe biological impacts. Determination of spatial extent and development of information relevant to pollutant source control at these sites may require additional investigation. A number of other sites demonstrated elevated chemical concentrations without severe acute toxicity, and still other sites had toxic sediment without having elevated concentrations of measured chemicals. These sites may warrant further studies of chronic effects and/or investigations to determine the likely causes of observed biological impacts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AA	Atomic Absorption
ASTM	American Society for Testing Materials
BPTCP	Bay Protection and Toxic Cleanup Program
CDFG	California Department of Fish and Game
CH	Chlorinated Hydrocarbon
COC	Chain of Custody
COR	Chain of Records
EDTA	Ethylenediaminetetraacetic Acid
EMAP	Environmental Monitoring and Assessment Program
EPA	Environmental Protection Agency
EqP	Equilibrium Partitioning
ERL	Effects Range - Low
ERM	Effects Range - Median
ERMQ	Effects Range - Median, Quotient
FAAS	Flame Atomic Absorption Spectroscopy
GC/ECD	Gas Chromatograph Electron Capture Detection
GFAAS	Graphite Furnace Atomic Absorption Spectroscopy
HCl	Hydrochloric Acid
HDPE	High-density Polyethylene
HMW PAH	High Molecular Weight Polycyclic Aromatic Hydrocarbons
HNO_3	Nitric Acid
HPLC/SEC	High Performance Liquid Chromatography Size Exclusion
H_2S	Hydrogen Sulfide
IDORG	Identification and Organizational Number
KCl	Potassium Chloride
LC_{50}	Lethal Concentration (to 50 percent of test organisms)
LMW PAH	Low Molecular Weight Polycyclic Aromatic Hydrocarbons
MDL	Method Detection Limit
MDS	Multi-Dimensional Scaling
MLML	Moss Landing Marine Laboratories
MPSL	Marine Pollution Studies Laboratory
MSD	Minimum Significant Difference
NH_3	Ammonia (Un-ionized)
NOAA	National Oceanic and Atmospheric Administration
NOEC	No Observed Effect Concentration
NS&T	National Status and Trends Program
PAH	Polycyclic Aromatic Hydrocarbons
PCB	Polychlorinated Biphenyl

LIST OF ABBREVIATIONS (CONTINUED)

PEL	Probable Effects Level
PELQ	Probable Effects Level Summary Quotient
PPE	Porous Polyethylene
PVC	Polyvinyl Chloride
QA	Quality Assurance
QAPP	Quality Assurance Project Plan
QC	Quality Control
REF	Reference
SCCWRP	Southern California Coastal Waters Research Project
SEM-AVS	Simultaneously Extracted Metals minus Acid Volatile Sulfide
SFBRWQCB	San Francisco Bay Regional Water Quality Control Board
SJSUF	San Jose State University Foundation
SPARC	Scientific Planning and Review Committee
SQC	Sediment Quality Criteria
SWRCB	State Water Resources Control Board
T	Temperature
TBT	Tributyltin
TEL	Threshold Effects Level
TFE	Tefzel Teflon®
TIE	Toxicity Identification Evaluation
TOC	Total Organic Carbon
TOF	Trace Organics Facility
UCSC	University of California at Santa Cruz
U.S. EPA	U.S. Environmental Protection Agency
WCS	Whole core squeezing

Units

1 part per thousand (ppt) = 1 mg/g sediment, or 1 g/L

1 part per million (ppm) = 1 mg/kg, 1 µg/g sediment, or 1 mg/L

1 part per billion (ppb) = 1 µg/kg, 1 ng/g sediment, or 1 µg/L

INTRODUCTION

Goals and Objectives

Legislative Mandate

This report presents and interprets data to assist in the identification of toxic hot spots in San Francisco Bay. The sediment quality assessment information described here is in support of management activities to protect the quality of waters and sediments of the State from discharges of waste, in-place sediment pollution, and any other factor that can impact beneficial uses of enclosed bays, estuaries and coastal waters (pursuant to Sections 13390 et seq., 13140 and 13143 of the California Water Code).

In 1989, the California State legislature added to and modified the California Water Code, Division 7, Chapter 5.6 to establish the Bay Protection and Toxic Cleanup Program (BPTCP), a comprehensive program to protect the existing and future beneficial uses of California's enclosed bays and estuaries. State Senate and Assembly bills SB 475 (1989), SB 1845 (1990), AB 41 (1989), and SB 1084 (1993) modified Chapter 5.6 [Bay Protection and Toxic Cleanup (Water Code Sections 13390-13396.5)] Division 7 of the Water Code for this purpose. The BPTCP has provided a new focus on regional efforts to control pollution of the State's bays and estuaries by establishing a program to identify toxic hot spots and plan for their cleanup.

Program Goals and Objectives

The BPTCP has four major goals:

1) To provide protection of present and future beneficial uses of the bay and estuarine waters of California;
2) To identify and characterize toxic hot spots;
3) To plan for toxic hot spot cleanup or other remedial or mitigation actions; and
4) To develop prevention and control strategies for toxic pollutants that will prevent creation of new toxic hot spots or the perpetuation of existing ones within the bays and estuaries of the State.

San Francisco Bay Study Goals and Objectives

As part of the legislative mandate, the BPTCP has implemented regional monitoring studies to identify toxic hot spots (Water Code Section 13392.5). The four objectives of BPTCP regional monitoring are:

1) To identify locations in enclosed bays, estuaries, or the ocean that are potential or candidate toxic hot spots;
2) To determine the extent of biological impacts in portions of enclosed bays and estuaries not previously sampled (areas of unknown condition);
3) To confirm the extent of biological impacts in enclosed bays and estuaries that have been previously sampled; and
4) To assess the relationship between toxic pollutants and biological effects.

Definition of a Toxic Hotspot

Section 13391.5 of the Water Code defines toxic hot spots as: "...[L]ocations in enclosed bays, estuaries, or adjacent waters..., the pollution or contamination of which affects the interests of the State, and where hazardous substances have accumulated in the water or sediment to levels which (1) may pose a substantial present or potential hazard to aquatic life, wildlife, fisheries, or human health, or (2) may adversely affect the beneficial uses of the bay, estuary, or ocean waters as defined in the water quality control plans, or (3) exceeds adopted water quality or sediment quality objectives."

Although the Water Code provides some direction in defining a toxic hot spot, the definition presented in Section 13391.5 is broad, and has been further refined by the State Water Resources Control Board (SWRCB) and Regional Water Quality Control Boards (RWQCBs). For a candidate toxic hot spot to be considered a "known toxic hotspot," the RWQCB and SWRCB must approve that designation, and the site must be adopted into the consolidated statewide toxic hot spot cleanup plan. At that point the site shall be considered a known toxic hot spot and all the requirements of the Water Code shall apply to that site. A "candidate toxic hotspot" is a site that meets the following conditions, but has not yet been approved by the RWQCB and SWRCB.

A site meeting any one or more of the following conditions is considered to be a "candidate" toxic hot spot:

1. The site exceeds water or sediment quality objectives for toxic pollutants that are contained in appropriate water quality control plans, or exceeds water quality criteria promulgated by the U.S. Environmental Protection Agency (U.S. EPA).

2. The water or sediment exhibits recurrent toxicity associated with toxic pollutants that is significantly different from the toxicity observed at reference sites (*i.e.*, when compared to the lower confidence interval of the reference envelope), based on toxicity tests acceptable to the SWRCB or the RWQCBs. Toxic pollutants should be present in the media at concentrations sufficient to cause or contribute to toxic responses in order to satisfy this condition.

3. The tissue toxic pollutant levels of organisms collected from the site exceed levels established by the United States Food and Drug Administration (FDA) for the protection of human health, or the National Academy of Sciences (NAS) for the protection of human health or wildlife. When a health advisory against the consumption of edible resident non-migratory organisms has been issued by the Office of Environmental Health Hazard Assessment (OEHHA) or Department of Health Services (DHS), on a site or water body, the site or water body is automatically classified as a "candidate" toxic hot spot if the chemical contaminant is associated with sediment or water at the site or water body.

4. Impairment measured in the environment is associated with toxic pollutants found in resident individuals. Impairment means reduction in growth, reduction in reproductive capacity, abnormal development, histopathological abnormalities.

5. Significant degradation in biological populations and/or communities is associated with the presence of elevated levels of toxic pollutants.

A "potential toxic hot spot" is defined as a suspect site with existing information indicating possible impairment but without sufficient information to be classified further as a candidate toxic hot spot. Additional details of the toxic hot spot definition used in the BPTCP can be found in SWRCB (1994).

Studies Designed to Meet Program Objectives

A phased process using effects-based measurements was used to identify toxic hot spots in California's enclosed bays and estuaries. In the San Francisco Bay Region, this process included: 1) a review of approximately 100 reports on water and sediment quality; 2) a survey of numerous bay and wetland sites in the Pilot Regional Monitoring Program (PRMP), which also included a methods validation study along a pollution gradient; 3) a reference site study that evaluated toxicological and statistical methods for identifying polluted sites by comparison with reference conditions, 4) a toxicity screening study of 127 stations from throughout the region (selected on the basis of previous information and PRMP results); and 5) confirmation studies of sites that exhibited toxicity and/or elevated chemistry during screening. The preliminary studies are described below. The screening and confirmation studies were designed to incorporate three measures, commonly known as the sediment quality triad. The triad approach consists of toxicity testing, benthic community analysis, and chemical analysis for trace metals and organic compounds. An additional bioaccumulation component was employed at selected sites.

While toxicity testing was used for screening the majority of stations, stations could also be evaluated with benthic community analyses or chemical tests or bioaccumulation data to provide sufficient information to list a site as a potential toxic hot spot or a site of concern. Sediment grain size, total organic carbon (TOC), NH_3 and H_2S concentrations were measured on all samples. While these factors are naturally occurring components of benthic environments, TOC, NH_3 and H_2S concentrations can be enhanced by human activities. They were quantified to assist in the interpretation of biological analyses.

A positive result or an observed adverse effect in any of the triad tests would trigger the confirmation step (depending on available funding). The confirmation phase consisted of performing all components of the sediment quality triad: toxicity, benthic community analysis, and chemical analysis, on the previously sampled site of concern. Assessment of benthic community structure was not included at all confirmation stations due to difficulty in interpreting the information for parts of San Francisco Bay, as described in greater detail in the Methods Section. Sediment samples from many of the confirmation stations were analyzed in laboratory bioaccumulation tests to determine whether sediment-associated chemicals were bioavailable to exposed biota.

Study Area

Geography, Hydrology and Biology

The San Francisco Bay Region is comprised of most of the San Francisco estuary up to the mouth of the Sacramento-San Joaquin Delta (Figure 1). The San Francisco estuary conveys the waters of the Sacramento and San Joaquin rivers into the Pacific Ocean. Located on the central coast of California, the Bay system functions as the only drainage outlet for waters of the Central Valley. It also marks a natural topographic separation between the northern and southern coastal mountain ranges.

The Sacramento and San Joaquin rivers, which enter the Bay system through the Delta at the eastern end of Suisun Bay, contribute almost all of the freshwater inflow to the Bay. Many smaller rivers and streams also convey fresh water to the Bay system. The rate and timing of these freshwater flows are among the most important factors influencing physical, chemical and biological conditions in the estuary. Flows in the region are highly seasonal, with more than 90 percent of the annual runoff occurring during the winter rainy season between November and April.

San Francisco Bay is typical of estuaries worldwide in that it provides critical habitat for aquatic species, including many commercially and ecologically important marine species that use estuaries as rearing grounds for sensitive early life-stages (Conomos et al., 1979). San Francisco Bay is also home to hundreds of introduced exotic species, brought in over the last 150 years, primarily in ship ballast water. The San Francisco estuary is made up of many different types of aquatic habitats that support a great diversity of organisms. Suisun Marsh in Suisun Bay is the largest brackish-water marsh in the United States. San Pablo Bay is a shallow embayment strongly influenced by runoff from the Sacramento and San Joaquin Rivers. The Central Bay is the portion of the Bay most influenced by oceanic conditions. The South Bay, with less freshwater inflow than the other portions of the Bay, acts more like a tidal lagoon. Together these areas sustain rich communities of aquatic life and serve as important wintering sites for migrating waterfowl and spawning areas for anadromous fish.

Human Uses of Land and Waterways

The natural harbor provided by the Bay has made it the focus of tremendous economic and industrial activity. There are three major port facilities, located at Oakland, San Francisco, and Richmond, as well as numerous tanker moorings associated with local refining activities, and many past and presently utilized military docks and shipyards. Most industrial sectors are represented along the Bay shores, including metal works, chemical manufacturing plants, ship yards, oil refineries, military bases, commercial fishing and shipping facilities, salt evaporation ponds, agriculture, and construction of residential and commercial buildings to serve the fourth largest metropolitan area in the United States. The Bay is at the center of an urban area including all or major portions of 9 counties: Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma.

General Sources of Pollution

San Francisco Bay is an extremely dynamic depositional environment. Sediments flow from the major river systems and storm water channels and are deposited in the Bay. Strong winds and tidal currents resuspend and redeposit these sediments, resulting in a system where sediments are well mixed. Bioaccumulative contaminants attach to sediments and are distributed and mixed by the same physical processes. Chemical transport is complex, and the sediments act as a sink for pollutants. These sediments may then act as sources of pollutants to organisms in the aquatic food chain and ultimately to humans.

The San Francisco estuary has high concentrations of metals due to contributions from numerous sources, both natural and anthropogenic. Natural sources include runoff from geologic formations, such as the local Franciscan Formation and the distant Sierra Nevada mountains and foothills. These formations are naturally enriched in some metals, including nickel, chromium, and mercury. Localized deposits of these metals were unearthed in a great wave of mining activity from the 1820's, continuing, in some cases, into the 1970's. Mercury was mined at numerous locations in the Coastal Range and then transported to the Sierra Nevada foothills to be used in the amalgamation of gold in placer and hydraulic mining. Drainage from natural mercury deposits, mine tailings, and processing activities, both in the coast range and the Sierra Nevada, is part of a complex transport process leading to elevated concentrations of mercury in Bay sediments and organisms. Transport of some naturally occurring metals, such as selenium, has been enhanced by leaching from irrigated agricultural soils. Selenium is also a waste product of oil refining activities. Metals such as copper, zinc, and silver are components of industrial and municipal wastewater discharged to the Bay, and industrial slag deposits may be responsible for some locally elevated concentrations of these and other metals.

Organic chemicals enter the Bay from a variety of sources. PCBs have accumulated in the sediments of the estuary due to past use and deposition. PCB mixtures were used extensively in the U.S. prior to 1979, when their manufacture, processing, use and application was banned, except in totally enclosed applications such as transformers. PCBs were used for industrial applications requiring fluids with thermal stability, low flammability, oxidation resistance, and solubility in organic compounds, and their widespread use provided a variety of opportunities for transport into the Bay. PCBs have proven to be extremely persistent in the environment. Regional Monitoring Program (RMP) data indicate that PCBs exceed non-promulgated U.S. EPA water quality criteria in water-column samples from throughout the estuary. This is probably due to resuspension from Bay sediments. BPTCP monitoring has shown that, except for a few areas described in this report, PCBs are fairly well mixed in the sediments of the estuary, which act as an ongoing source to organisms up the food chain. Storm events can mobilize PCBs from soils and urban surfaces and transport them into the estuary. Recent monitoring by the RMP has indicated that Coyote Creek may be a current source of PCB loads into the South Bay (SFEI, 1997). Increased monitoring is necessary to identify and cleanup any ongoing sources.

Pesticide transport to the Bay is greatly affected by compound class and history of use. Persistent organochlorine pesticides, such as the DDTs, chlordanes and dieldrin that were identified in the BPTCP fish study (SFBRWQCB, 1995), have similar properties in that they are

extremely persistent in the environment and highly lipid soluble. Since these lipid soluble compounds are not easily metabolized or excreted, they are stored in fatty tissue and can readily bioaccumulate in fish tissue with high lipid content. Although these chemicals (and many other organochlorines) have been banned for use in the U.S. for approximately 20 years, they persist in soils that wash into the Bay, and are still commonly detected at elevated concentrations in sediments and tissue from throughout the estuary. One large historic source of DDT, Lauritzen Channel in Richmond Harbor, has been recently dredged and capped. Other sources may be detected through increased monitoring of sediment-laden stormwater.

The newer generation organophosphate, carbamate and pyrethroid pesticides are more water-soluble and accumulate to a lesser degree in sediments and tissues, yet they may be responsible for biological impacts to water-column species. Many of these compounds are widely available for residential use, and have been found in toxic concentrations in urban stormwater entering the Bay (Hansen et al., 1993). They are also applied in large quantities on farmlands, and have been implicated in toxicity of Central Valley agricultural drain water (Norberg-King et al., 1991). The Central Valley RWQCB has supported numerous studies investigating the transport and impacts of these compounds in the Delta, tributary rivers, and urban creeks.

Polycyclic aromatic hydrocarbons (PAHs) are a group of compounds that vary considerably in toxicity and are widely distributed in Bay sediments. Low-molecular-weight PAHs tend to be more acutely toxic, while high-molecular-weight PAHs have greater carcinogenic potential, often after being metabolized in organisms (Kennish, 1998). Sources of PAHs include oil, municipal and industrial wastewaters, combustion of fossil fuels, and urban runoff.

As industrial and municipal waste treatment has improved, and population around the Bay has increased, urban stormwater runoff has become the major source for mass loading of pollutants that accumulate in the food chain and of pesticides that cause acute toxicity to aquatic organisms. In the past several years, the RMP and the Bay Area Stormwater Management Agencies Association (BASMAA) have been conducting limited monitoring studies of runoff from urban creeks. Through this monitoring, Coyote Creek has been identified as a source of PCBs and chlorinated pesticides to the estuary. In other urban creeks, high levels of toxicity have been identified during runoff events, possibly due to the pesticides diazinon and chlorpyrifos (e.g., Hansen et al., 1993). Identification of the sources of these contaminants and the development of watershed management plans are necessary to protect the beneficial uses of the estuary. U.S. EPA and the State Board strongly encourage the development of watershed management plans to protect watersheds draining to the Bay. However, there must be increased watershed monitoring and assessment in order to identify and prioritize current or potential problems, so that watershed management plans can be adequately targeted and evaluated.

Previous and Concurrent Monitoring Programs

Pilot Regional Monitoring Program

In addition to the screening and confirmation of toxic hot spots, several other studies have been conducted through the BPTCP in this region. In 1991 and 1992, the Pilot Regional Monitoring Program (PRMP) was conducted. The main purpose of this study was to develop

the design and methodology for an ongoing regional monitoring program. The PRMP also had a screening component where sediment chemistry and toxicity was measured in wetlands throughout the Bay. The third component was a gradient study, conducted in Castro Cove, to validate methods for the BPTCP and RMP.

Regional Monitoring Program

In 1993, the San Francisco Estuary Regional Monitoring Program for Trace Substances (RMP) was established. The program is administered through the San Francisco Estuary Institute and funded by municipal and industrial entities that discharge wastewater to the Bay, and by those involved in bay dredging activities. Through this program, chemistry and toxicity data are collected from water-column and sediment samples collected at established sites along the central spine of the estuary. Bioaccumulation is measured in tissues of transplanted bivalves from throughout the estuary, and the program conducts a number of special studies each year investigating sources and causes of observed biological effects. The data collected has been valuable in identifying seasonal, temporal and spatial trends in contamination, toxicity, and benthic community structure, as well as in the investigation of physical and biological processes affecting chemical transport and exposure. These data are available in annual reports (e.g., SFEI, 1997) and through the SFEI website, newsletters and other media.

Fish Tissue Study

In 1994, a study was conducted under the BPTCP to measure contaminant levels in fish from San Francisco Bay (SFBRWQCB, 1995). This was the first study conducted in the Bay to investigate whether contaminant concentrations were elevated in fish being consumed by the public, and to determine whether a health advisory was necessary. Results of the study indicated that six chemicals or chemical groups were of potential concern. These chemicals were mercury, PCBs, DDT, dieldrin, chlordane and dioxins. As a result of the study, the Office of Environmental Health Hazard Assessment (OEHHA) issued an interim health advisory on consuming fish caught in San Francisco Bay and the Delta. Regular monitoring of contaminants in fish, studies on consumption patterns, public outreach, and education projects are currently being conducted in this Region to address these concerns.

Sediment Reference Site Study

In 1994 and 1995, a study was conducted to identify sediment reference sites in San Francisco Bay, to identify toxicity test methods that would be most appropriate for use in the Bay, and to develop a statistical method to distinguish between sites representing ambient conditions and those where locally elevated chemical concentrations may have been responsible for biological impacts (Hunt et al., 1998). This study was necessary because varying levels of sediment toxicity had been observed even in less polluted sites that were relatively remote from sources, and standard statistical methods could not adequately distinguish this level of toxicity from more severe levels observed at other sites. Since the purpose of the BPTCP was to identify toxic hot spots, new methods needed to be developed that could distinguish between ambient conditions, as represented by a distribution of reference site data, and more toxic sites that might be considered for remedial action. This study identified five reference sites in the Bay (two in San Pablo Bay, one in the Central Bay and two in the South Bay), evaluated nine different toxicity tests for use in toxic hot spot

screening and confirmation studies, and developed a statistical method to distinguish between ambient conditions and potential toxic hot spots. Once reference sites were identified, toxicity tests chosen, and the statistical method developed, screening and confirmation studies began.

Methods Validation: Effects of Extraction Method on Porewater Toxicity

As part of the development and validation of analytical methods for the BPTCP, a study was conducted in 1993 to assess the effects of three extraction methods on the toxicity of sediment porewater. Sediment was collected at four sites in the San Francisco Bay Region, and porewater was extracted by centrifugation, piston squeezing, and passive separation through sediment settlement. The results of this study are reported in Appendix G of this report.

Studies of Additional Sites of Concern

There are additional sites of concern in the San Francisco Bay Region that do not technically qualify as candidate toxic hot spots under the definition used in this program. Most of these sites are redevelopment properties or military bases slated for closure. Many of these sites are the subject of large-scale investigations, including environmental risk assessments. Lauritzen Channel, which was previously listed as a potential toxic hot spot in 1993, was investigated under CERCLA; this site was dredged and the bottom then capped with sand during the summer of 1997.

At military bases, sediment pollution is evaluated in the larger context of determining the risk to human and ecological receptors in ecological risk assessments required under the federal CERCLA program. CERCLA is the primary regulatory authority driving environmental investigations at military bases, and these investigations are generally extensive. Jurisdictions other than the Regional Board, including the U.S. EPA, the U.S. Fish and Wildlife Service, the National Oceanic and Atmospheric Administration, the California Department of Fish and Game and the California Department of Toxic Substances Control also participate in designing and determining the scope of these characterizations. Although some of these sites were visited by the BPTCP early in the program, and efforts were made by others to use methods and protocols consistent with those of the BPTCP, the study designs used and the scale of these investigations were distinctly different from the BPTCP studies reported here. Because of funding constraints and a desire to avoid regulatory overlap, evaluations of these sites was left to responsible agencies and not further addressed in the BPTCP.

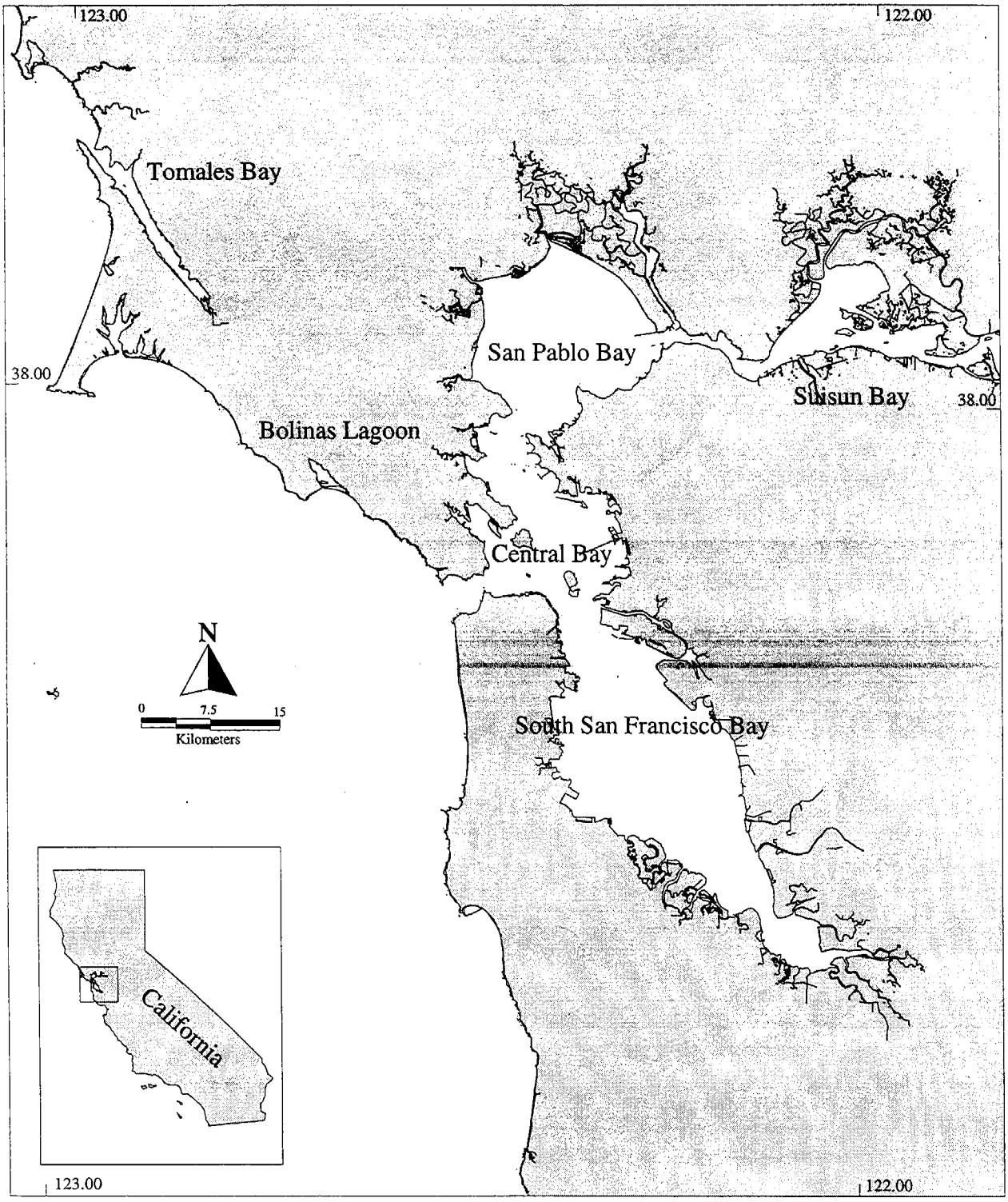

Figure 1. Location of Region 2 Sampling Areas.

METHODS

Study Approach

Sampling Design and Analyses Conducted

Rationale

Sampling site selection was based on knowledge of geographic and hydrologic characteristics of the Bay, sources of chemical inputs, and results of previous sediment assessment studies. Sites were selected with the specific intention of investigating those areas thought most likely to be affected by pollutants. Randomized sampling was not incorporated into the study design. The primary sources for previous information used to identify screening sites included the Pilot Regional Monitoring Program described by Flegal et al. (1994), , the sediment toxicity studies of Long et al. (1990), State Mussel Watch data (e.g., Phillips, 1988) and approximately 100 additional reports on water and sediment quality that were reviewed by RWQCB staff.

Screening and Confirmation Sampling and Analysis

Due to the size of the San Francisco Bay, the large number of potential problem sites, and funding limitations, it was not possible to fully characterize every site where pollution problems might exist. In order to focus efforts on the sites of greatest concern, a two-phase sampling process was used. In the first phase, 127 sites that had been identified through previous investigations were screened for sediment toxicity (Figure 2). While toxicity testing may not detect chronically toxic compounds and does not directly predict impacts to biological communities (Luoma and Carter, 1993), it was used as the screening tool because the tests are simple, precise indicators of the integrated effects of sediment contaminants (Swartz et al., 1985). Knowledge of sediment chemistry alone is currently insufficient to predict biological effects (Chapman et al., 1987), and measures of benthic community structure in San Francisco Bay can be confounded by salinity fluctuations and the impact of invading exotic species (Nichols and Thompson, 1985). Toxicity testing was therefore selected as the primary screening tool.

After reviewing the screening data and information from previous studies, a number of stations were resampled as part of the confirmation phase of the study (Figure 3). Twelve stations were resampled and analyzed with the sediment quality triad, including two toxicity tests and sediment chemical analysis at all confirmation stations, and benthic community analysis at multiple stations within seven main confirmation sites. Ten of these confirmation stations were also analyzed for bioaccumulation using 28-day laboratory exposures with the clam *Macoma nasuta*. A total of 46 stations were screened for a broad suite of trace metal and organic compounds, and a total of 143 samples were analyzed for mercury and PCBs, chemicals that were identified as elevated in fish tissues in the Bay (Fairey et al., 1997). An additional 15 stations were resampled and tested with sea urchin larvae in sediment-water interface exposures, because their screening samples exhibited toxicity only in porewater tests that were accompanied by elevated sulfide or ammonia levels.

Sediment reference sites identified in a previous study (Hunt et al., 1998) were sampled during each screening and confirmation survey. Toxicity data from these reference sites were added to the reference data base from which toxicity tolerance limits were calculated (see Statistical Analyses section, below). Maps of sampling stations are presented in Figure 2. Specific station location information, including GIS latitude and longitude data, are provided in Appendix B.

There was no field replication in either the screening or confirmation phases of this study. Consideration of spatial extent of pollution was generally deferred to future studies of sites identified in screening and confirmation. However, linear transects leading away from some confirmation stations provided opportunities to investigate relationships between sediment chemistry and biological effects along contamination gradients. Phase I sediment Toxicity Identification Evaluations (TIEs) were conducted at two sites, and an abbreviated sediment-water interface TIE was conducted at a third site to investigate possible causes of sediment toxicity.

Surficial sediments were collected in this program to evaluate the most biologically active layer of sediment and its effects on aquatic organisms. Due to the dynamic nature of the sediments in this Region, sediment samples were collected to a depth of 5 cm, the same depth that is sampled in the RMP. In other BPTCP regional studies, the sample depth was 2 cm.

Sample Collection and Processing

Summary of Methods

Specific techniques used for collecting and processing samples are described in this section. Because collection of sediments influences the results of all subsequent laboratory and data analyses, it was important that samples be collected in a consistent and conventionally acceptable manner. Field and laboratory technicians were trained to conduct a wide variety of activities using standardized protocols to ensure comparability in sample collection among crews and across geographic areas. Sampling protocols in the field followed the accepted procedures of EMAP, NS&T, and ASTM and included methods to avoid cross-contamination; methods to avoid contamination by the sampling activities, crew, and vessel; collection of representative samples of the target surficial sediments; careful temperature control, homogenization and subsampling; and chain of custody procedures.

Cleaning Procedures

All sampling equipment (i.e., containers, container liners, scoops, water collection bottles) was made from non-contaminating materials and was precleaned and packaged protectively prior to entering the field. Sample collection gear and samples were handled only by personnel wearing non-contaminating polyethylene gloves. All sample collection equipment (excluding the sediment grab) was cleaned by using the following sequential process: Two-day soak and wash in Micro® detergent, three tap-water rinses, three deionized water rinses, a three-day soak in 10% HCl, three ASTM Type II Milli-Q® water rinses, air dry, three petroleum ether rinses, and air dry.

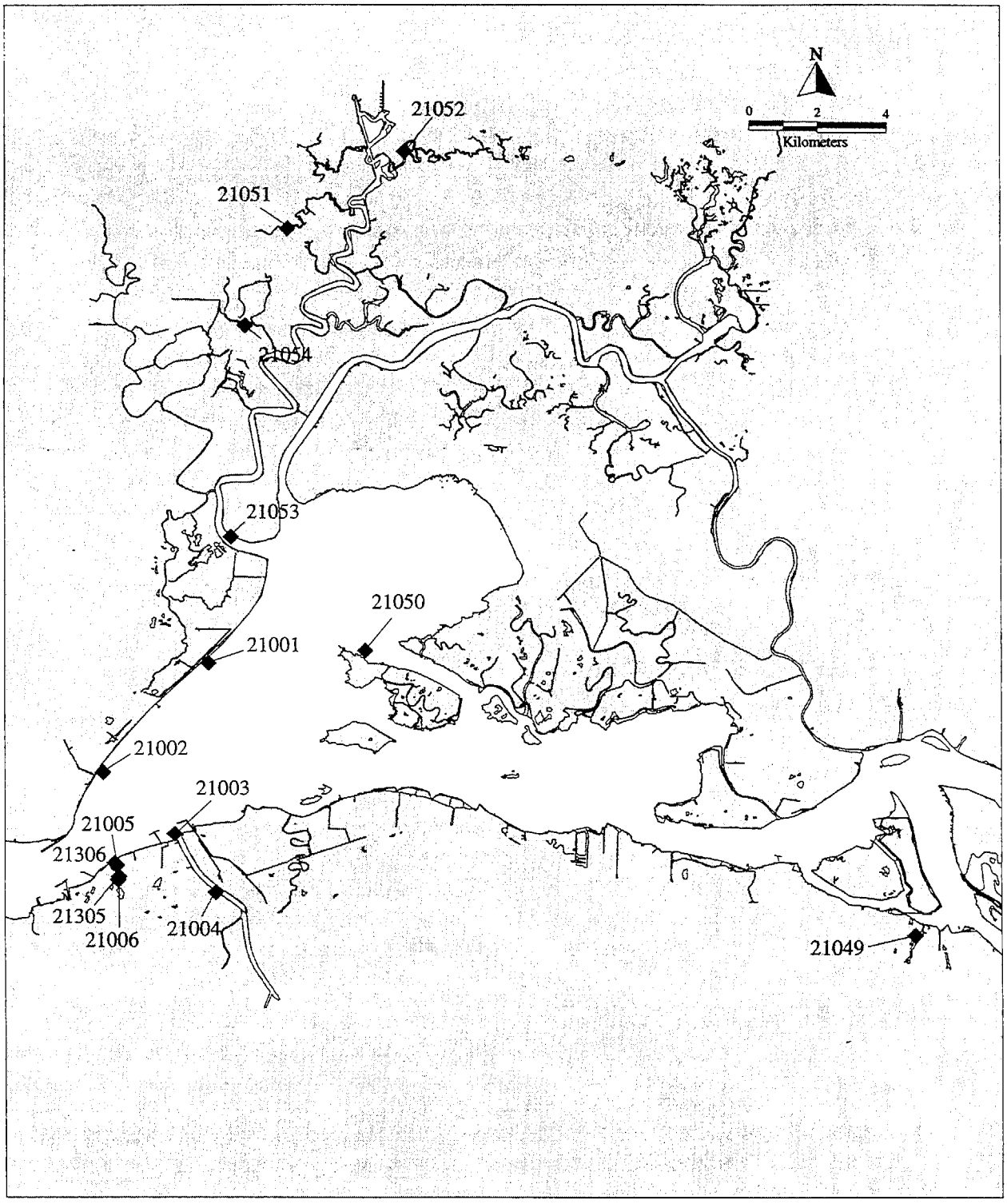

Figure 2a. Sampling Locations in Suisun Bay.

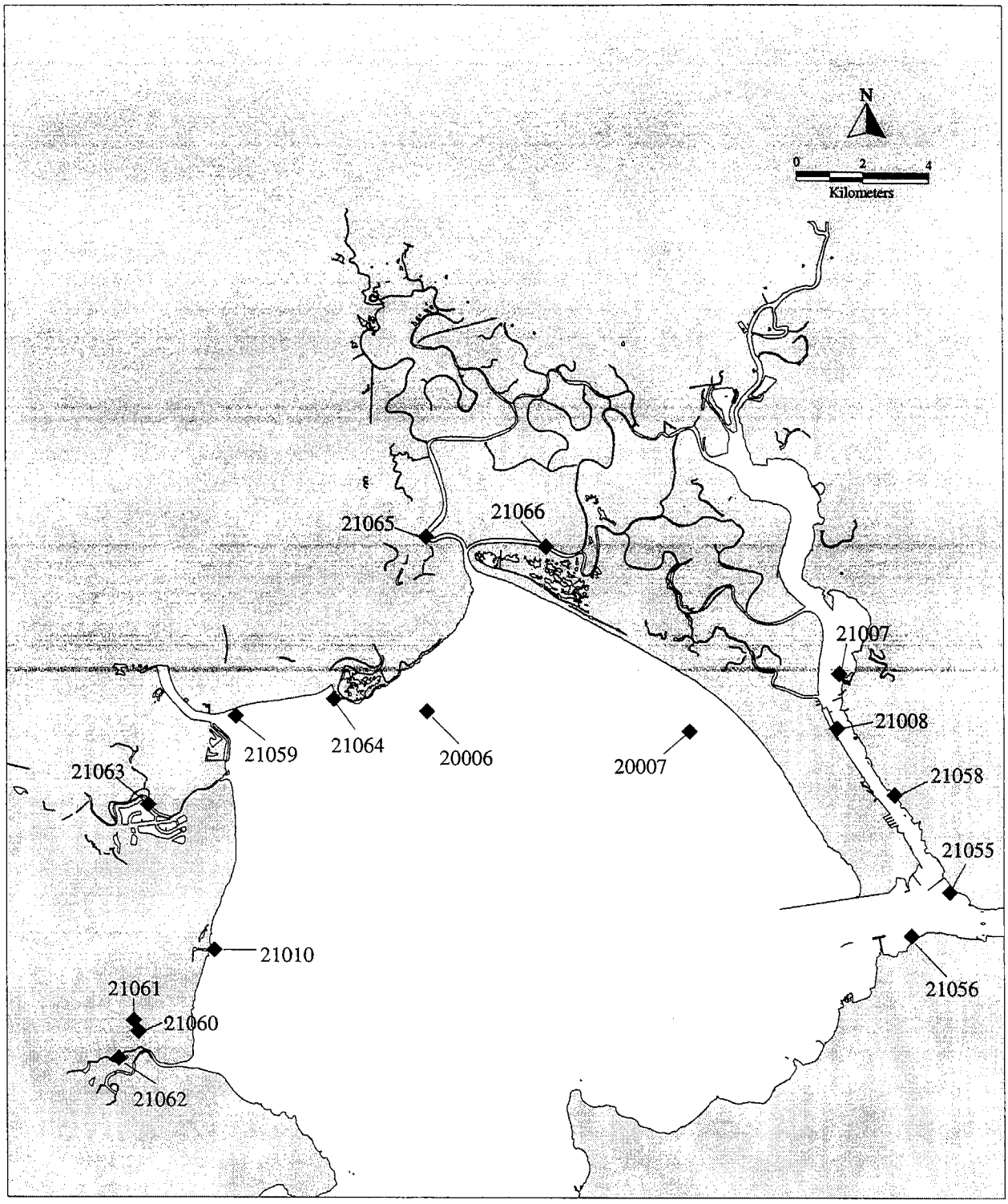

Figure 2b. Sampling Locations in San Pablo Bay.

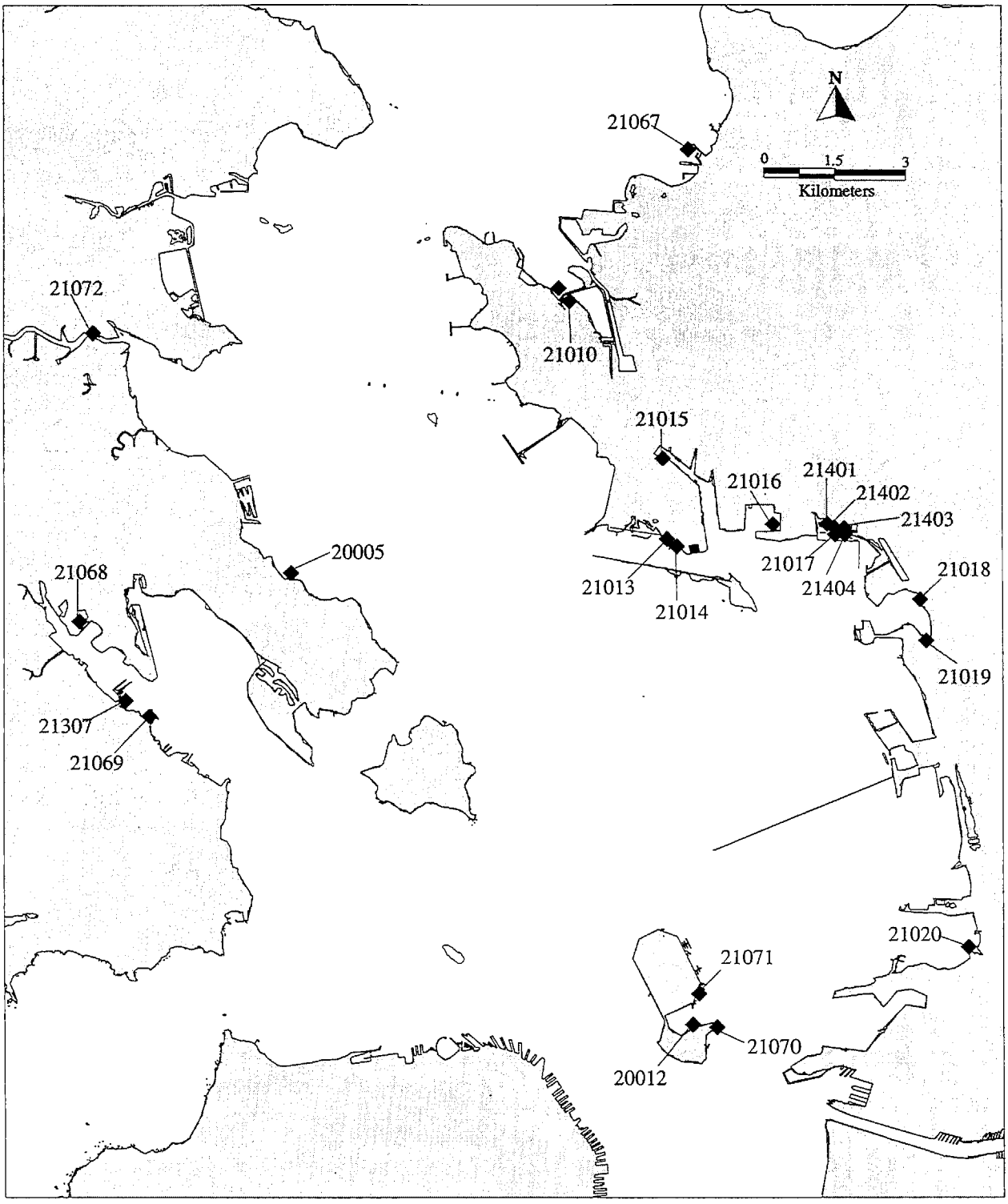

Figure 2c. Sampling Locations in Central San Francisco Bay.

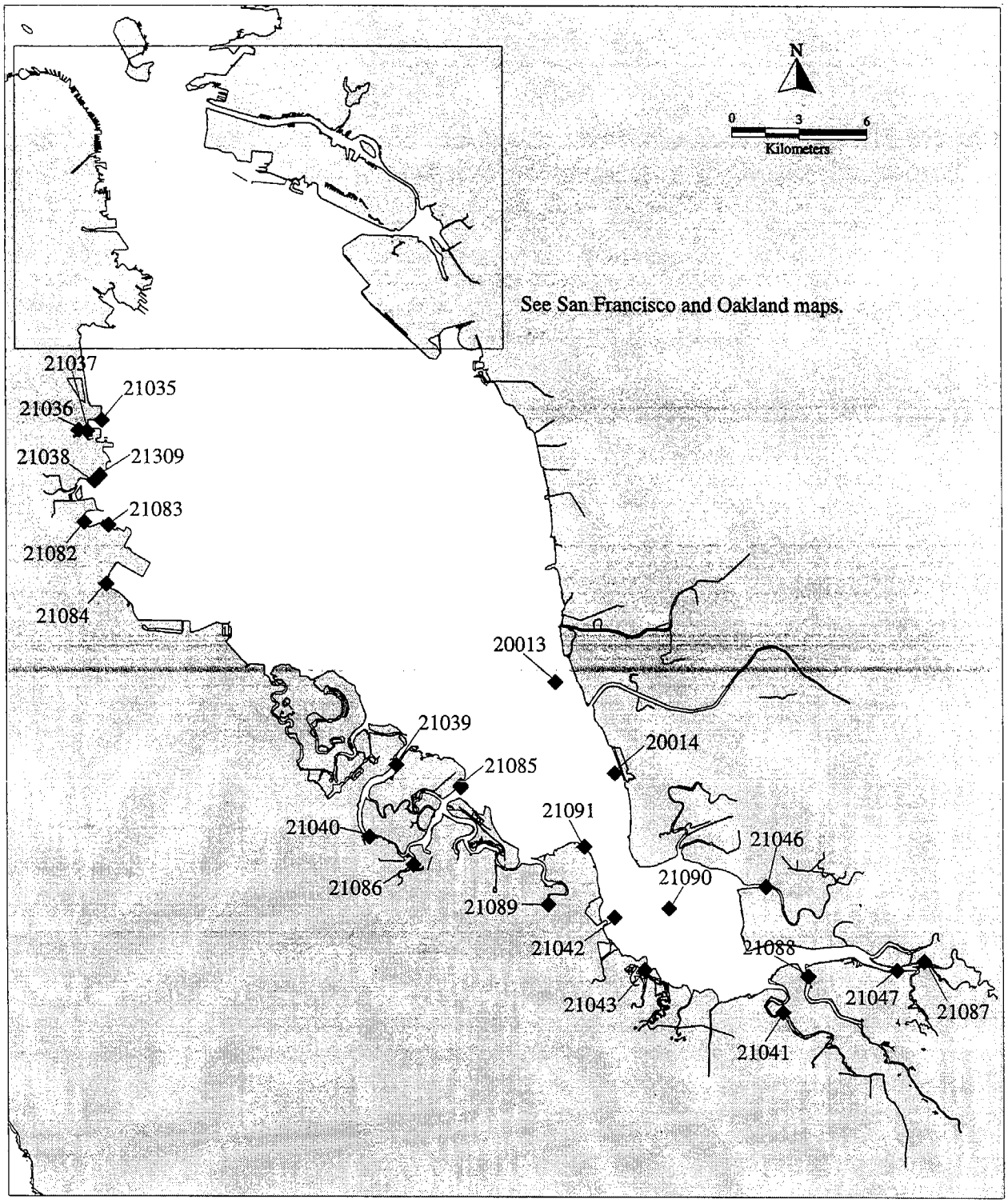

Figure 2d. Sampling Locations in South San Francisco Bay.

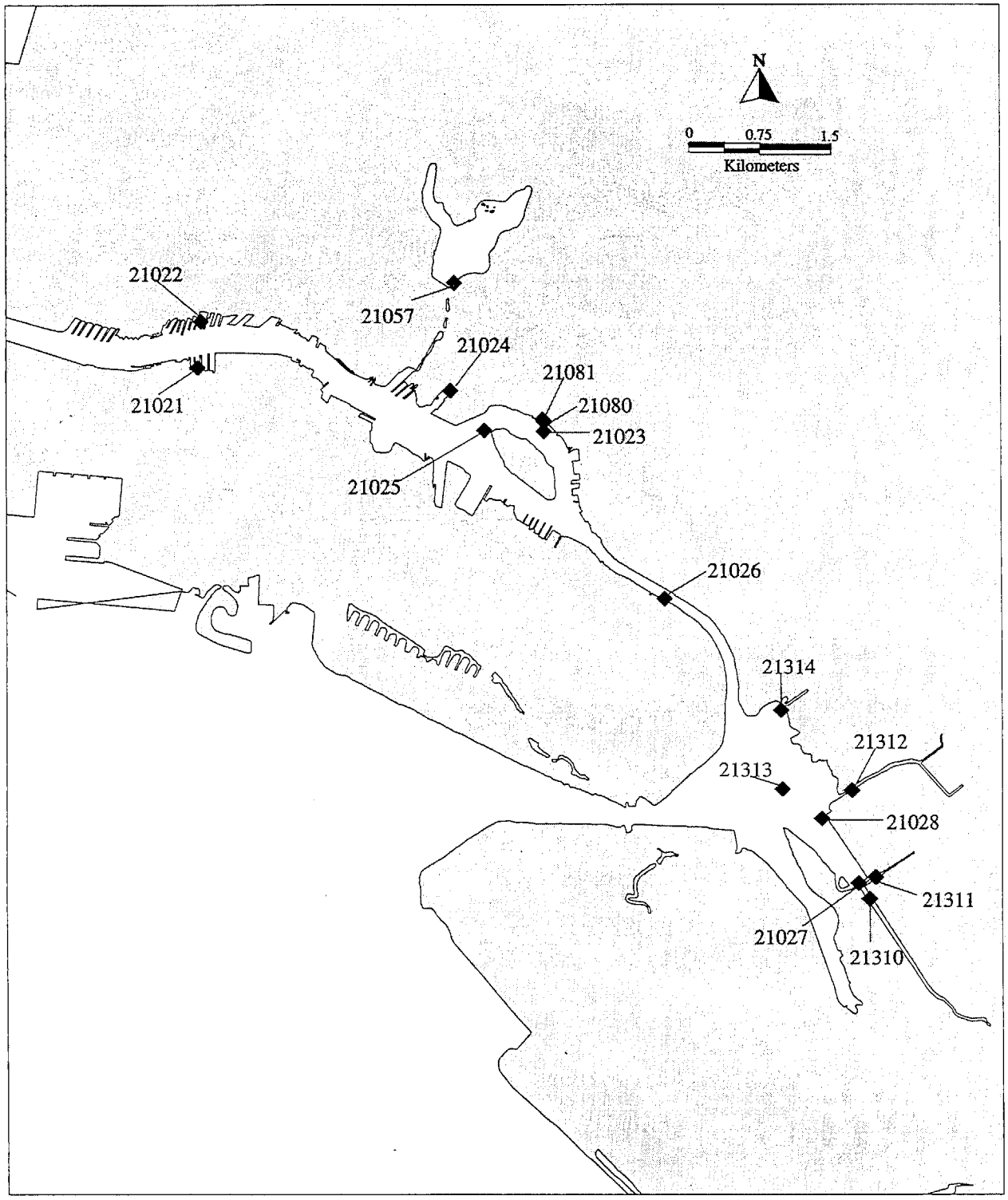

Figure 2e. Sampling Locations in Oakland.

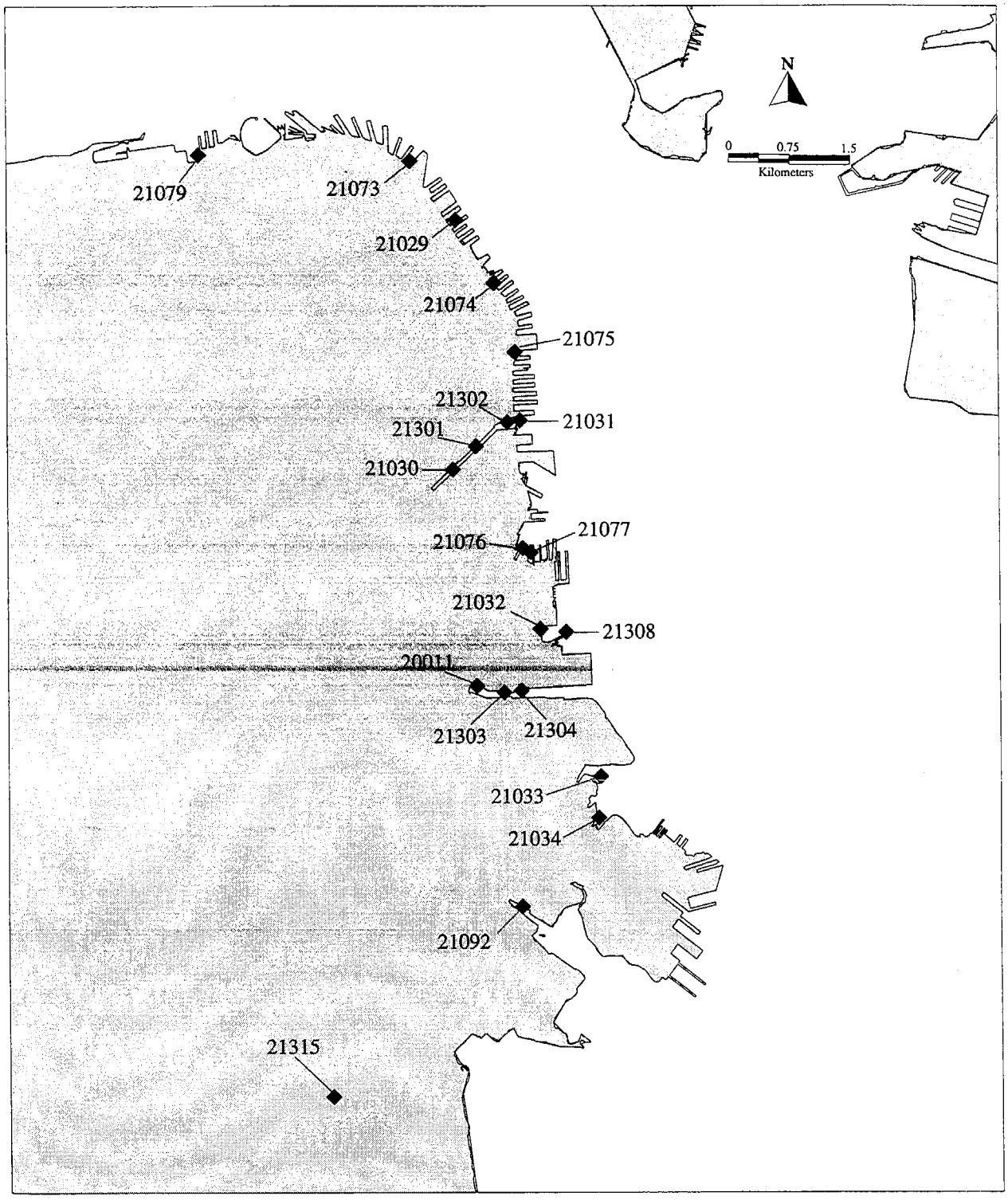

Figure 2f. Sampling Locations in San Francisco.

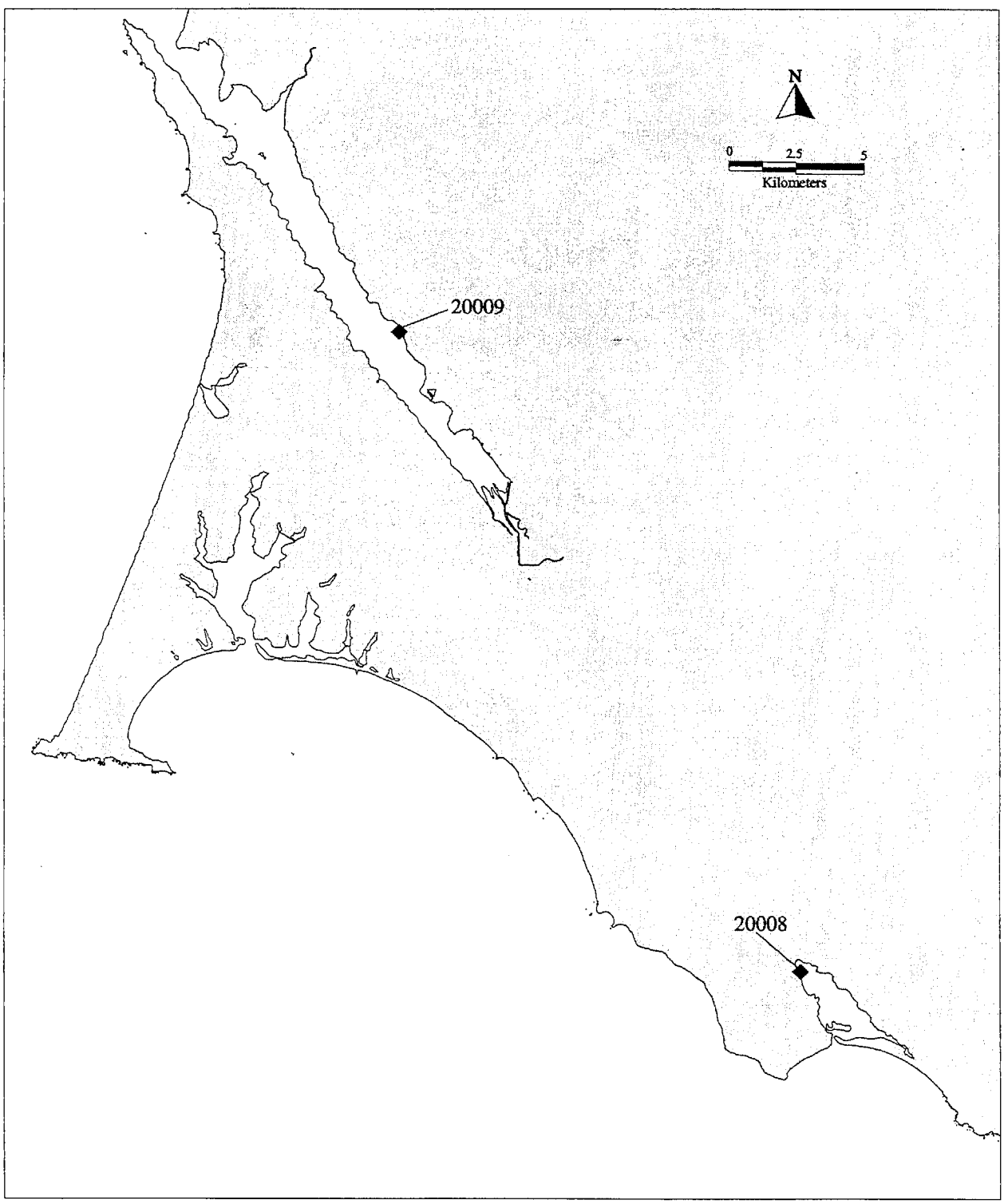

Figure 2g. Sampling Locations in Tomales Bay and Bolinas Lagoon.

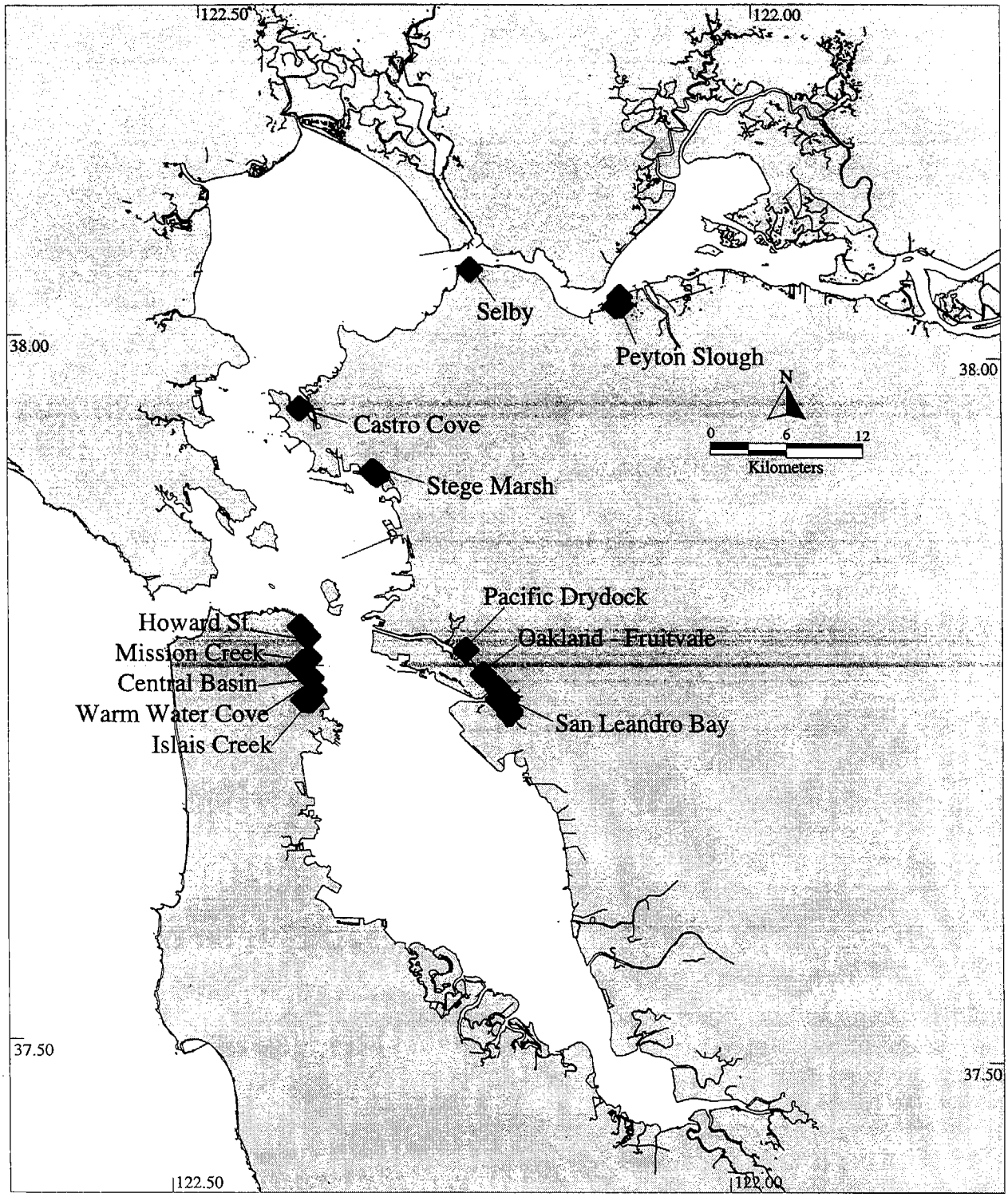

Figure 3. Location of Region 2 Confirmation Stations.

All cleaning after the Micro® detergent step was performed in a positive pressure "clean" room to prevent airborne contaminants from contacting sample collection equipment. Air supplied to the clean room was filtered.

The sediment grab was cleaned prior to entering the field, and between sampling stations, by utilizing the following sequential steps: a vigorous Micro® detergent wash and scrub, a sea-water rinse, a 10% HCl rinse, and a methanol rinse. The sediment grab was scrubbed with seawater between successive deployments at the same station to remove adhering sediments from contact surfaces possibly originating below the sampled layer.

Sample storage containers were cleaned in accordance with the type of analysis to be performed upon its contents. All containers were cleaned in a positive pressure "clean" room with filtered air to prevent airborne contaminants from contacting sample storage containers.

Plastic containers (HDPE or TFE) for trace metal analysis media (sediment, archive sediment, pore water, and subsurface water) were cleaned by: a two-day Micro® detergent soak, three tap-water rinses, three deionized water rinses, a three-day soak in 10% HCl or HNO_3, three Type II Milli-Q® water rinses, and air dry.

Glass containers for total organic carbon, grain size or synthetic organic analysis media (sediment, archive sediment, pore water, and subsurface water) and additional Teflon sheeting cap-liners were cleaned by: a two-day Micro® detergent soak, three tap-water rinses, three deionized water rinses, a three-day soak in 10% HCl or HNO_3, three Type II Milli-Q® water rinses, air dry, three petroleum ether rinses, and air dry.

Sediment Sample Collection

All sampling locations (latitude & longitude), whether altered in the field or predetermined, were verified using a Magellan NAV 5000 Global Positioning System, and recorded in the field logbook. The primary method of sediment collection was by use of a $0.1m^2$ Young-modified Van Veen grab aboard a sampling vessel. Modifications include a non-contaminating Kynar coating which covered the grab's sample box and jaws. After the filled grab sampler was secured on the boat gunnel, the sediment sample was inspected carefully. The following acceptability criteria were met prior to taking sediment samples. If a sample did not meet all the criteria, it was rejected and another sample was collected.

1. Grab sampler was not over-filled (*i.e.*, the sediment surface was not pressed against the top of the grab).
2. Overlying water was present, indicating minimal leakage.
3. Overlying water was not excessively turbid, indicating minimal sample disturbance.
4. Sediment surface was relatively flat, indicating minimal sample disturbance.
5. Sediment sample was not washed out due to an obstruction in the sampler jaws.
6. Desired penetration depth was achieved (*i.e.*, 10 cm).
7. Sample was muddy (>30% fines), not sandy or gravelly.
8. Sample did not include excessive shell, organic or man-made debris.

It was critical that sample contamination be avoided during sample collection. All sampling equipment (*i.e.*, siphon hoses, scoops, containers) was made of non-contaminating material and was cleaned appropriately before use. Samples were not touched with un-gloved fingers. In addition, potential airborne contamination (*e.g.*, from engine exhaust, cigarette smoke) was avoided. Before sub-samples from the grab sampler were taken, the overlying water was removed by slightly opening the sampler, being careful to minimize disturbance or loss of fine-grained surficial sediment. Once overlying water was removed, the top 5 cm of surficial sediment was sub-sampled from the grab. Subsamples were taken using a precleaned flat bottom scoop. This device allowed a relatively large sub-sample to be taken from a consistent depth. When subsampling surficial sediments, unrepresentative material (e.g., large stones or vegetative material) was removed from the sample in the field. Small rocks and other small foreign material remained in the sample. Determination of overall sample quality was determined by the chief scientist in the field. Such removals were noted on the field data sheet. For the sediment sample, the top 5 cm was removed from the grab and placed in a pre-labeled polycarbonate container. Between grabs or cores, the sediment sample in the container was covered with a Teflon sheet, and the container covered with a lid and kept cool. When a sufficient amount of sediment was collected, the sample was covered with a Teflon sheet assuring no air bubbles. A second, larger Teflon sheet was placed over the top of the container to ensure an air tight seal, and nitrogen was vented into the container to purge it of oxygen.

If water depth did not permit boat entrance to a site (*e.g.*,<1 meter), divers sampled that site using sediment cores (diver cores). Cores consisted of a 10 cm diameter polycarbonate tube, 30 cm in length, including plastic end caps to aid in transport. Divers entered a study site from one end and sampled in one direction, so as to not disturb the sediment with feet or fins. Cores were taken to a depth of at least 15 cm. Sediment was extruded out of the top end of the core to the prescribed depth of 5-cm, removed with a polycarbonate spatula and deposited into a cleaned polycarbonate tub. Additional samples were taken with the same seawater rinsed core tube until the required total sample volume was attained. Diver core samples were treated the same as grab samples, with Teflon sheets covering the sample and nitrogen purging. All sample acceptability criteria were met as with the grab sampler.

Replicate benthic samples (n = 3) were obtained at predetermined sites from separate deployments of the sampler. The coring device was 10 cm in diameter and 14 cm in height, enclosing a 0.0075 m^2 area. Corers were placed into sediment with minimum disruption of the surface sediments, capturing essentially all surface-active fauna as well as species living deeper in the sediment. Corers were pushed about 12 cm into the sediment and retrieved by digging along one side, removing the corer and placing the intact sediment core into a PVC screening device. Sediment cores were sieved through a 0.5 mm screen and residues (*e.g.*, organisms and remaining sediments) were rinsed into pre-labeled storage bags and preserved with a 10% formalin solution. After 3 to 4 days, samples were rinsed and transferred into 70% isopropyl alcohol and stored for future taxonomy and enumeration.

Transportation of Samples

Six-liter sample containers were packed (three to an ice chest) with enough ice to keep them cool for 48 hours. Each container was sealed in precleaned, large plastic bags closed with a cable tie to prevent contact with other samples or ice or water. Ice chests were driven back to the laboratory by

the sampling crew or flown by air freight within 24 hours of collection.

Homogenization and Aliquoting of Samples
Samples remained in ice chests (on ice, in double-wrapped plastic bags) until the containers were brought back to the laboratory for homogenization. All sample identification information (station numbers, etc.) was recorded on Chain of Custody (COC) and Chain of Record (COR) forms prior to homogenizing and aliquoting. A single container was placed on plastic sheeting while also remaining in original plastic bags. The sample was stirred with a polycarbonate stirring rod until mud appeared homogeneous.

All prelabeled jars were filled using a clean Teflon or polycarbonate scoop and stored in freezer/refrigerator (according to media/analysis) until analysis. The sediment sample was aliquoted into appropriate containers for trace metal analysis, organic analysis, pore water extraction, and bioassay testing. Samples were placed in boxes sorted by analysis type and leg number. Sample containers for sediment bioassays were placed in a refrigerator (4°C) while sample containers for sediment chemistry (metals, organics, TOC and grain size) were stored in a freezer (-20°C).

Procedures for the Extraction of Pore water
Pore water was extracted using centrifugation. All pore water extraction procedures were performed using trace metal and trace organic clean techniques in a positive pressure clean room with filtered air to prevent airborne contamination.

All sample containers or sampling equipment in contact with sediment or pore water received a scrub and 2 day soak in MICRO® detergent, followed by triple fresh and deionized water rinses. Equipment was then immersed in 10% HCl for 3 days, triple rinsed in MILLI-Q® Type II water, air dried, and triple rinsed with petroleum ether.

Samples were stored on ice at 4°C prior to centrifugation. Pre-cleaned Teflon scoops were used to transfer sediment from sample containers into high-speed one-liter polycarbonate centrifuge jars, which were spun at 2500 G for 30 minutes at 4°C in a Beckman J-6B refrigerated centrifuge.

Porewater was transferred from each centrifuge jar into final sample containers (250 pre-cleaned borosilicate glass jars) using pre-cleaned polyethylene siphons. While decanting, care was taken to avoid floating debris, fauna, shell fragments or other solid material. After transfer into final sample containers, porewater was immediately refrigerated at 4°C. Samples were refrigerated, not frozen, and testing was initiated within 24 hours of extraction of the final samples.

Chain of Records & Custody
Chain-of-records documents were maintained for each station. Each form was a record of all sub-samples taken from each sample. IDORG (a unique identification number for only that sample), station numbers and station names, leg number (sample collection trip batch number), and date collected were included on each sheet. A Chain-of-Custody form accompanied every sample so that each person releasing or receiving a subsample signed and dated the form.

Authorization/Instructions to Process Samples

Standardized forms entitled "Authorization/Instructions to Process Samples" accompanied the receipt of any samples by any participating laboratory. These forms were completed by DFG personnel, or its authorized designee, and were signed and accepted by both the DFG authorized staff and the staff accepting samples on behalf of the particular laboratory. The forms contain all pertinent information necessary for the laboratory to process the samples, such as the exact type and number of tests to run, number of laboratory replicates, dilutions, exact eligible cost, deliverable products (including hard and soft copy specifications and formats), filenames for soft copy files, expected date of submission of deliverable products to DFG, and other information specific to the lab/analyses being performed.

Toxicity Testing

Summary of Methods

All toxicity tests were conducted at the California Department of Fish and Game (DFG) Marine Pollution Studies Laboratory (MPSL) at Granite Canyon. Toxicity tests were conducted by personnel from the Institute of Marine Sciences, University of California, Santa Cruz. The following toxicity tests were conducted in this study: infaunal amphipod *Eohaustorius estuarius* 10-day survival in solid-phase sediment, sea urchin *Strongylocentrotus purpuratus* 96-hour embryo-larval development in sediment porewater, and sea urchin *Strongylocentrotus purpuratus* 96-hour embryo-larval development in sediment-water interface (SWI) exposures to intact cores of solid-phase sediment. Two freshwater samples were each tested with the amphipod *Hyalella azteca* 10-day survival test in solid-phase sediment and the water flea *Ceriodaphnia dubia* 96-hour survival test in sediment porewater.

Sediment Samples

Bedded sediment samples were transported to MPSL from the sample-processing laboratory at Moss Landing in ice chests at 4°C. Transport time was approximately one hour. Samples were held at 4°C, and all tests were initiated within 14 days of sample collection, unless otherwise noted in the Quality Assurance Appendix. All sediment samples were handled according to procedures described in ASTM (1992) and BPTCP Quality Assurance Project Plan (Stephenson et al. 1993). Samples were removed from refrigeration the day before the test, and loaded into test containers. Water quality was measured at the beginning and end of all tests. At these times pH, temperature, salinity, and dissolved oxygen were measured in overlying water from all samples to verify that water quality criteria were within the limits defined for each test protocol. Total ammonia concentrations were also measured at these times. Samples of overlying and interstitial water for hydrogen sulfide measurement were taken at the beginning and end of each toxicity test. Hydrogen sulfide samples were preserved with zinc acetate and stored in the dark until time of measurement.

Porewater Samples

Once at MPSL, pore water samples were stored in the dark, at 4°C. Samples were equilibrated to test temperature (15°C) on the day of the test, and pH, temperature, salinity, and dissolved oxygen were measured in all samples to verify water quality criteria were within the limits defined for the test protocol. Total ammonia and sulfide concentrations were also measured. Pore water samples with salinities outside specified ranges for each protocol were adjusted to within the acceptable

range. Salinities were increased by the addition of hypersaline brine, which was drawn from partially frozen seawater at a salinity of 60 to 80‰. In cases where original sample salinity was very low, addition of hypersaline brine diluted the samples to as low as 55% of their original strength, thus similarly diluting any potential toxins present. Dilution was greatest in North Bay and marsh samples during winter surveys (sampling dates are given in the appendices). Water quality parameters, as mentioned above, were measured at the beginning and end of each pore water test.

Measurement of Ammonia and Hydrogen Sulfide

Total ammonia concentrations were measured using an Orion Model 95-12 Ammonia Electrode. The concentration of unionized ammonia was derived from the sample pH and concentration of total ammonia using the following equation (from Whitfield 1974, 1978):

$$[NH_3] = [\text{total ammonia}] \times ((1 + \text{antilog}(pK_a° - pH))^{-1}),$$

where $pK_a°$ is the stoichiometric acidic hydrolysis constant for the test temperature and salinity. Values for $pK_a°$ were experimentally derived by Khoo *et al.* (1977). The method detection limit for total ammonia was 0.1 mg/L.

Total sulfide concentrations were measured on a spectrophotometer using a colorimetric method (Phillips et al. 1997). The concentration of hydrogen sulfide was derived from the concentration of total sulfide by using the following equation (ASCE 1989):

$$[H_2S] = [S^{2-}] \times (1 - ((1 + \text{antilog}(pK_a° - pH))^{-1})),$$

where temperature and salinity dependent $pK_a°$ values were taken from Savenko (1977). The method detection limit for total sulfide was 0.01 mg/L. Values and corresponding detection limits for unionized ammonia and hydrogen sulfide were an order of magnitude lower than those for total ammonia and total sulfide, respectively. Care was taken with all sulfide and ammonia samples to minimize volatilization by keeping water quality sample containers capped tightly until analysis. Initial and final ammonia and sulfide values were measured, and the highest of these two values were reported for each sample tested.

Effects of Unionized Ammonia and Hydrogen Sulfide

Toxicity results were screened against known application limits for unionized ammonia and hydrogen sulfide (Table 1). Toxicity test ammonia and sulfide concentrations above the application limits were taken into consideration when examining toxicity test results.

Table 1. Unionized ammonia and hydrogen sulfide effects thresholds for BPTCP toxicity test protocols.

Species	Unionized Ammonia (mg/L)	Limit Definition	Reference
Ampelisca	0.4	Application Limit	EPA 1995
Eohaustorius	0.8	Application Limit	EPA 1995
Red Abalone	0.05	NOEC	MPSL
Mytilus	0.15	LOEC	Tang et al. 1997
Neanthes	1.25	LOEC	Dillon 1993
Rhepoxynius	0.4	Application Limit	EPA 1995
Purple Urchin Development	0.07	NOEC	Bay et al. 1993
Purple Urchin Fertilization	>1.4	NOEC	Bay et al. 1993

Species	Hydrogen Sulfide (mg/L)	Limit Definition	Reference
Eohaustorius	0.114	LOEC	Knezovich et al. 1996
Mytilus	0.0053	LOEC	Knezovich et al. 1996
Rhepoxynius	0.087	LOEC	Knezovich et al. 1996
Purple Urchin Development	0.0076	LOEC	Knezovich et al. 1996
Purple Urchin Fertilization	0.007-0.014	NOEC	Bay et al. 1993

Amphipod (*Eohaustorius estuarius*) Survival Tests

Solid-phase sediment toxicity was assessed in 127 samples using the 10-day amphipod survival toxicity test protocols outlined in EPA 1994. All *Eohaustorius* test amphipods were obtained from Northwestern Aquatic Sciences (NWAS) in Yaquina Bay, Oregon. Animals were separated into groups of approximately 100 and placed in polyethylene boxes containing Yaquina Bay collection site sediment, then shipped on ice via overnight courier. Upon arrival at Granite Canyon, the *Eohaustorius* were acclimated to 20‰ (T=15°C). Once acclimated, the animals were held for an additional 48-hours prior to addition to the test containers.

Test containers were one liter glass beakers or jars containing 2 cm of sediment and filled to the 700-ml line with control seawater adjusted to the appropriate salinity using spring water or distilled well water. Test sediments were not sieved for resident organisms prior to testing although at the conclusion of the test, the presence of any predators was noted and recorded on the data sheet. Test sediment and overlying water were allowed to equilibrate for 24 hours, after which 20 amphipods were placed in each beaker along with control seawater to fill test containers to the one-liter line. Test chambers were aerated gently and illuminated continuously at ambient laboratory light levels.

Five laboratory replicates of each sample were tested for ten days. A negative sediment control consisting of five lab replicates Yaquina Bay home sediment was included with each sediment test. After ten days, the sediments were sieved through a 0.5-mm Nitex screen to recover the test animals. The number of survivors was recorded for each replicate, and percent survival was calculated as:

$$\frac{\text{(Number of surviving amphipods)}}{\text{(Initial number of amphipods)}} \times 100$$

Positive control reference tests were conducted concurrently with each sediment test using cadmium chloride as a reference toxicant. For these tests, amphipod survival was recorded in three replicates of four cadmium concentrations after a 96-hour water-only exposure. Negative controls for the reference test consisted of one micron-filtered Granite Canyon seawater, diluted to the appropriate salinity.

Sea Urchin (*Strongylocentrotus purpuratus*) Embryo-Larval Development Test

The sea urchin (*Strongylocentrotus purpuratus*) larval development test was conducted on 95 pore water samples. Details of the test protocol are given in EPA 1995. A brief description of the method follows.

Sea urchins were collected from the Monterey County coast near Granite Canyon, and held at MPSL at ambient seawater temperature and salinity (33±2‰) until testing. Adult sea urchins were held in complete darkness to preserve gonadal condition. On the day of a test, urchins were induced to spawn in air by injection with 0.5M KCl. Eggs and sperm collected from the urchins were mixed in seawater at a 500 to 1 sperm to egg ratio, and embryos were distributed to test containers within 1 hour of fertilization. Test containers were polyethylene-capped, seawater leached, 20-ml glass scintillation vials containing 10 mLs of sample. Each test container was inoculated with approximately 250 embryos (25/ml). Forty seven of the 95 pore water samples were tested at three concentrations: 100, 50 and 25% pore water, each having three replicates. These samples were diluted with one micron-filtered Granite Canyon seawater. Laboratory controls were included with each set of samples tested. Controls include a dilution water control consisting of Granite Canyon seawater, and a brine control with all samples that require brine adjustment. Tests were conducted at ambient seawater salinity (33±2‰). A 96-hour positive control reference test was conducted concurrently with each pore water test using a dilution series of copper chloride as a reference toxicant.

After a 96-hour exposure, larvae were fixed in 5% buffered formalin. Approximately 100 larvae in each container were examined under an inverted light microscope at 100x to determine the proportion of normally developed larvae as described in EPA 1995. Visual clues used to identify embryos as normal included development of skeletal rods (spicules) that extend beyond half the length of the larvae and normal development of a three-part gut. Embryos demonstrating retarded development were considered abnormal. Percent normal development was calculated as:

$$\frac{\text{Number of normally developed larvae counted}}{\text{Total number of larvae counted}} \times 100$$

Sea Urchin Test Using The Sediment-Water Interface Exposure System

The purple sea urchin (*Strongylocentrotus purpuratus*) embryo/larval development test at the sediment-water interface was conducted on 40 intact core sediment samples taken with minimal disturbance from the Van Veen grab sampler or directly from in-place sediments. The method follows Anderson et al. (1996); a brief description follows.

Sea urchins were collected, handled, spawned and fertilized as described above. Each sediment-water interface test container consisted of a polycarbonate tube with a 25-µm screened bottom placed inside the sediment core tube so that the screen was within 1 cm of the sediment surface. Seawater at ambient salinity was gently poured into the core tube and allowed to equilibrate for 24 hours before the start of the test. After inserting the screen tube into the equilibrated cores, each tube was inoculated with approximately 250 embryos. The laboratory control consisted of similar core tubes holding Yaquina Bay amphipod home sediment provided by NWAS. Tests were conducted at ambient seawater salinity ± 2‰. Ambient salinity at Granite Canyon is usually 32 to 34‰. A water-only positive control reference test was conducted concurrently with the test using a dilution series of copper chloride as a reference toxicant.

After an exposure period of 96 hours, larvae were fixed in 5% buffered formalin. One hundred larvae in each container were examined under an inverted light microscope at 100x to determine the proportion of normally developed larvae as described in EPA 1995. Percent normal development was calculated as:

$$\frac{\text{Number of normally developed larvae counted}}{\text{Total number of larvae counted}} \times 100$$

Water Flea (*Ceriodaphnia dubia*) Acute Survival Test

Aquatic toxicity of two freshwater samples was assessed using the Cladoceran water flea *Ceriodaphnia dubia* acute survival test. The method follows EPA (1993); a brief description follows.

Ceriodaphnia neonates (<24 h) were obtained either from in-house cultures or from Toxscan Laboratories (Watsonville, CA). Neonates were isolated on Day 0 of the test. All dilution water was prepared according to EPA (1993). Pore water test containers were 50-mL glass beakers containing 15 mL of test solution. Each test container was inoculated with 5 or 8 neonates depending on availability. The laboratory negative control consisted of EPA dilution water. After an exposure period of 96 hours neonates were counted. A positive control reference test was conducted concurrently with the test using a dilution series of copper chloride as the reference toxicant.

Freshwater Amphipod (*Hyalella azteca*) Survival Test

Freshwater amphipod tests were conducted on two samples following the ASTM (1993) *Hyalella azteca* protocol. All *Hyalella* were obtained from NWAS in Yaquina Bay, Oregon. Animals were separated into groups of approximately 1000 and placed in polyethylene cubitainers containing NWAS laboratory water, then shipped via overnight courier. Upon arrival at Granite Canyon, the amphipods were acclimated to Granite Canyon well water (25°C). Once acclimated, the animals were held for an additional 48-h prior to addition to the test containers.

Test containers were one-liter glass jars containing 2 cm of sediment and filled to the 700-mL line with Granite Canyon well water. Test sediment and overlying water were allowed to equilibrate for 24 hours, then 20 amphipods were placed in each beaker along with well water to fill each test container to the one-liter line. Test chambers were gently aerated and continuously illuminated.

Five replicates of each sample were tested for 10 days. In addition, a negative sediment control consisting of 5 replicates of Yaquina Bay home sediment was included with each set of samples tested. Test containers were fed a slurry of crushed alfalfa pellets three times per week (ASTM 1993). After 10 days, samples were sieved through a 0.5-mm Nitex screen to recover the test animals, and the number of survivors was recorded for each replicate. Percent survival was calculated as:

$$\frac{\text{(Number of surviving amphipods)}}{\text{(Initial number of amphipods)}} \times 100$$

Positive control reference tests were conducted concurrently with each sediment test using cadmium chloride as a reference toxicant. In these tests, amphipod mortality was recorded in three replicates of four cadmium concentrations after a 96-hour water-only exposure. A dilution water control consisting of Granite Canyon well water was included in each test.

Test Acceptability and Evaluation

Quality Assurance/Quality Control (QA/QC) guidelines for the toxicity tests used in the BPTCP project are summarized in the BPTCP Quality Assurance Project Plan (Stephenson, et al. 1993). Test acceptability criteria from published protocols were evaluated for all tests. Quality assurance checklists were compiled that noted compliance for all tests with each of these criteria. Evaluation codes were assigned to each deviation from QA/QC guidelines, and can be summarized as follows:

-3: sample has minor exceedances of QA criteria that are unlikely to affect assessments.
-4: sample meets or exceeds control criteria requirements.
-5: data has exceedances, but are generally usable for most assessments and reporting purposes.
-6: sample has major exceedances of control criteria requirements and the data is not usable for most assessments and reporting purposes. No toxicity test data in this report warranted this designation.

It is recommended that the QA evaluations listed in the appendices be consulted before using any data reported here for critical management decisions.

Toxicity Identification Evaluations (TIEs)

TIEs with Strongylocentrotus purpuratus

Phase I TIEs are designed to characterize samples by isolating broad classes of compounds to determine their relationship to observed toxicity. Phase I TIE procedures include adjustment of sample pH, chelation of cationic compounds (including many trace metals), neutralization of oxidants (such as chlorine), aeration to remove volatiles, inactivation of metabolically activated toxicants, solid-phase extraction (SPE) of non-polar organic compounds on C-18 columns and subsequent elution of extracted compounds. Each sample fraction, in which classes of

compounds have been removed, inactivated, or isolated, is then tested for toxicity. TIE procedures followed the methods described by US EPA (1996). These procedures are described briefly below. All TIE treatments were tested with the purple urchin embryo-larval development test, as described above. Treatment solution (sample fraction) was divided into 5 replicate 20-mL scintillation vials (10 mL of solution), with approximately 250 embryos inoculated into each vial. Each sample was tested at three dilutions. The sample underwent the TIE treatment prior to being diluted with one micron-filtered Granite Canyon seawater that had also undergone the TIE treatment. Testing sample dilutions provides information on the degree of sample toxicity. The TIE treatments are described as follows:

Aeration
Sample was aerated for one hour to remove volatile compounds.

Filtration
Sample was filtered through a 0.45 μm glass fiber filter to remove toxicants associated with particulate material.

Graduated pH
Adjusting sample pH can affect the toxicity of hydrolizable, ionic, acidic, or basic compounds. Sample pH was adjusted and maintained at pH 7.9, 8.1 and 8.4 by the addition of hydrochloric acid and/or sodium hydroxide.

EDTA Chelation
Addition of EDTA binds cationic trace metals, such as copper, cadmium, mercury, zinc, lead, nickel, and, too a lesser extent, silver and manganese, resulting in relatively non-toxic metal complexes (Hockett and Mount 1996). EDTA was added to the sample for a final concentration of 60 mg/L. The sample was allowed to interact with the EDTA for three hours before the pH was adjusted with sodium hydroxide. pH was checked prior to distributing sample into test containers.

Sodium Thiosulfate Addition
Addition of sodium thiosulfate (STS) reduces oxidants, such as chlorine, ozone, chlorine dioxide, mono- and di-chloroamines, bromine, iodide, manganous ions, and certain electrophylic organic chemicals (Mount and Anderson-Carnahan, 1988a). It also binds some trace metals, such as copper, cadmium, mercury, silver, and too a lesser extent, zinc, lead, and nickel (Hockett and Mount, 1996). STS was added to the sample for a final concentration of 50 mg/L. The sample was allowed to interact for one hour.

Solid Phase Extraction (SPE)
Solid-phase extraction through a C-18 SPE column was used to remove a range of non-polar organic compounds from sample solutions. The SPE columns were later eluted with 100% methanol to allow toxicity testing of compounds retained on the column. Sample was pumped through silicone tubing that had been cleaned by running 25 mL of distilled water followed by 25 mL of methanol through each tubing apparatus (but not through the column). The column was

prepared by pumping 30 mL of methanol through it, followed by 50 mL of distilled water. Next, laboratory dilution water was pumped through the column; the first 20 mL was discarded, and the remaining volume was kept as the column control solution. Finally, 200 mL of sample was run through the column; the first 20 mL was discarded, and the remaining volume collected as SPE treated sample. Column was kept wet until all sample had been passed through.

Column Eluate

After C-18 SPE extraction described above, the column was then run dry and air-dried with a syringe. With the stopcock tightly shut, 2 mL of 100% methanol was added to the column. The stopcock was then opened, and air pumped into the column at 2 mL/min until the column was dry. Eluate was collected in a small vial. The 2 mL aliquot of eluate was then delivered into 800 mL of laboratory dilution water. Assuming that all non-polar organic constituents from the sample were retained on the column (no breakthrough), and assuming that all of these compounds were then completely removed from the column in the methanol eluate, then the eluate treatment (2 mL in 800 mL) would contain 25% of the concentration of these constituents as did the original sample (25% addback). This low level of addback was necessitated by scarcity of pore water sample and by volume requirements for treatment testing. An eluate control consisting of 2 mL of methanol added to 800 mL of laboratory dilution water was tested with each C-18 eluate treatment.

Piperonyl Butoxide Tests

A number of organophosphate pesticides (phosphorothioate compounds such as diazinon, chlorpyrifos, malathion, parathion, methyl parathion and fenthion) require metabolic activation by exposed organisms before they become toxic. These activation reactions consist of oxidative metabolism by the cytochrome P-450 group of enzymes (Durhan et al. 1993). This activation can be blocked by compounds such piperonyl butoxide (PBO), thereby reducing or eliminating toxicity due to this class of compounds.

In this study, PBO was added to test samples to determine whether metabolically activated pesticides were responsible for observed toxicity. 2.5 mL of 50 mg/L PBO stock solution was added to 250 mL of each sample (resulting in a concentration of 0.5 mg/L PBO). PBO controls were made by adding 20 mL PBO to 180 mL of laboratory dilution water.

Abbreviated *Strongylocentrotus purpuratus* Sediment-Water Interface TIEs with EDTA

The purple sea urchin (*Strongylocentrotus purpuratus*) embryo/larval development test at the sediment-water interface (SWI) was conducted on intact core sediment samples taken with minimal disturbance. Details of the SWI test protocol are described above. Duplicate cores were collected for EDTA and blank treatments. Addition of EDTA binds cationic trace metals, such as copper, cadmium, mercury, zinc, lead, nickel, and, too a lesser extent, silver and manganese, resulting in relatively non-toxic metal complexes (Hockett and Mount 1996). EDTA was mixed into seawater at ambient salinity for a final concentration of 50 mg/L, and the pH was adjusted back to ambient pH using sodium hydroxide. This EDTA-spiked seawater was then added as overlying water in the SWI containers, while untreated seawater was added as overlying water to the untreated samples. Overlying water was allowed to equilibrate for 24 hours before the start of the test. After inserting the screen tube into the equilibrated cores, each tube was inoculated

with approximately 250 embryos. The laboratory controls consisted of Yaquina Bay amphipod home sediment from Northwestern Aquatic Sciences with and without EDTA in the overlying water.

Trace Metals Analysis of Sediments

Summary of Methods

Trace metals analyses were conducted at the California Department of Fish and Game's (CDFG) Trace Metals Facility at Moss Landing, CA. Table 2 indicates the trace metals analyzed and lists method detection limits for sediments. These methods were modifications of those described by Evans and Hanson (1993), as well as those developed by the CDFG (California Department of Fish and Game, 1990). Samples were selected for chemical analyses by SWRCB staff based on results from toxicity tests.

Sediment Digestion Procedures

One gram aliquot of sediment was placed in a pre-weighed Teflon vessel, and one ml concentrated 4:1 nitric:perchloric acid mixture was added. Vessels were capped and heated in a vented oven at 130° C for four hours. Three ml hydrofluoric acid were added to the vessel, recapped and returned to oven overnight. Twenty ml of 2.5% boric acid were added to the vessel and placed in oven for an additional 8 hours. Weights of Teflon vessel and solution were recorded, and solution was poured into 30 ml polyethylene bottles.

Tissues Digestion Procedures

A three gram aliquot of tissue was placed in a pre-weighed Teflon vessel, and three mls of concentrated 4:1 nitric:perchloric acid mixture was added. Samples then were capped and heated on hot plates for five hours. Caps were tightened and heated in a vented oven at 130°C for four hours. Samples were allowed to cool and 15 mls of Type II water was added to the vessels. The solution was then quantitatively transferred to a pre weighed 30 ml polyethylene (HDPE) bottle and taken up to a final weight of 20 g with Type II water.

Atomic Absorption Methods

Samples were analyzed by furnace AA on a Perkin-Elmer Zeeman 3030 Atomic Absorption Spectrophotometer, with an AS60 auto sampler, or a flame AA Perkin Elmer Model 2280. Samples, blanks, matrix modifiers, and standards were prepared using clean techniques inside a clean laboratory. ASTM Type II water and ultra clean chemicals were used for all standard preparations. All elements were analyzed with platforms for stabilization of temperatures. Matrix modifiers were used when components of the matrix interferes with adsorption. The matrix modifier was used for Sn, Sb and Pb. Continuing calibration check standards (CLC) were analyzed with each furnace sheet, and calibration curves were run with three concentrations after every 10 samples. Blanks and standard reference materials, MESS1, PACS, BCSS1 or 1646 were analyzed with each set of samples for sediments.

Method Detection Limits

Table 2. Dry Weight Trace Metal Minimum Detection Limits

Analytes[†]	*MDL, µg/g dry Sediment*	*MDL, µg/g dry Tissue*	*MDL, µg/L Water*
Silver	0.002	0.01	0.001
Aluminum	1	1	NA
Arsenic	0.1	0.25	0.1
Cadmium	0.002	0.01	0.002
Copper	0.003	0.1	0.04
Chromium	0.02	0.1	0.05
Iron	0.1	0.1	0.1
Mercury	0.03	0.03	NA
Manganese	0.05	0.05	NA
Nickel	0.1	0.1	0.1
Lead	0.03	0.1	0.01
Antimony	0.1	0.1	NA
Tin	0.02	0.02	NA
Selenium	0.1	0.1	NA
Zinc	0.05	0.05	0.02

Trace Organic Analysis of Sediments (PCBs, Pesticides, and PAHs)

Summary of Methods

Analytical sets of 12 samples were scheduled such that extraction and analysis would occur within a 40 day window. Methods employed by UCSC-TOF were modifications of those described by Sloan *et al.* (1993). Tables 3-7 indicate the pesticides, PCBs, and PAHs currently analyzed, and list method detection limits for sediments and tissues on a dry weight basis.

Sediment Extraction

Samples were removed from the freezer and allowed to thaw. A 10 gram sample of sediment was removed for chemical analysis and an independent 10 gram aliquot was removed for dry weight determinations. The dry weight sample was placed into a pre-weighed aluminum pan and dried at 110°C for 24 hours. The dried sample was reweighed to determine the sample's percent moisture. The analytical sample was extracted 3 times with methylene chloride in a 250 mL amber Boston round bottle on a modified rock tumbler. Prior to rolling, sodium sulfate, copper, and extraction surrogates were added to the bottle. Sodium sulfate dehydrates the sample allowing for efficient sediment extraction. Copper, which was activated with hydrochloric acid, complexes free sulfur in the sediment. After combining the three extraction aliquots, the extract was divided into two portions, one for chlorinated hydrocarbon (CH) analysis and the other for polycyclic aromatic hydrocarbon (PAH) analysis.

Tissue Extraction

Samples were removed from the freezer and allowed to thaw. A 5 gram sample of tissue was removed for chemical analysis and an independent 5 gram aliquot was removed for dry weight determinations. The dry weight sample was placed into a pre-weighed aluminum pan and dried at

110°C for 24 hours. The dried sample was reweighed to determine the sample's percent moisture. The analytical sample was extracted twice with methylene chloride using a Tekmar Tissumizer. Prior to extraction, sodium sulfate and extraction surrogates were added to the sample and methylene chloride.

The two extraction aliquots were combined and brought to 100 mL. A 25 mL aliquot was decanted through a Whatmann 12.5 cm #1 filter paper into a pre-weighed 50 mL flask for lipid weight determination. The filter was rinsed with ~15 mL of methylene chloride and the remaining solvent was removed by vacuum-rotary evaporation. The residue was dried for 2 hours at 110°C and the flask was re-weighed. The change in weight was taken as the total methylene chloride extractable mass. This weight then was used to calculate the samples "percent lipid".

Organic Analysis

The CH portion was eluted through a silica/alumina column, separating the analytes into two fractions. Fraction 1 (F1) was eluted with 1% methylene chloride in pentane and contained > 90% of p,p'-DDE and < 10% of p,p'-DDT. Fraction 2 (F2) analytes were eluted with 100% methylene chloride. The two fractions were exchanged into hexane and concentrated to 500 µL using a combination of rotary evaporation, controlled boiling on tube heaters, and dry nitrogen blow downs.

F1 and F2 fractions were analyzed on Hewlett-Packard 5890 Series gas chromatographs utilizing capillary columns and electron capture detection (GC/ECD). A single 2 µl splitless injection was directed onto two 60 m x 0.25 mm I.D. columns of different polarity (DB-17 & DB-5; J&W Scientific) using a glass Y-splitter to provide a two dimensional confirmation of each analyte. Analytes were quantified using internal standard methodologies. The extract's PAH portion was eluted through a silica/alumina column with methylene chloride. It then underwent additional cleanup using size-exclusion high performance liquid chromatography (HPLC/SEC). The collected PAH fraction was exchanged into hexane and concentrated to 250 µL in the same manner as the CH fractions.

Table 3: Dry Weight Minimum Detection Limits of Chlorinated Pesticides.

Analytes	*Database Abbreviation*	*MDL, ng/g dry Sediment*	*MDL, ng/g dry Tissue*	*MDL, ng/L Water*
Fraction #1 Analytes †				
Aldrin	ALDRIN	0.5	1.0	2.0
alpha-Chlordene	ACDEN	0.5	1.0	1.0
gamma-Chlordene	GCDEN	0.5	1.0	1.0
o,p'-DDE	OPDDE	1.0	3.0	1.0
o,p'-DDT	OPDDT	1.0	4.0	2.0
Heptachlor	HEPTACHLOR	0.5	1.0	2.0
Hexachlorobenzene	HCB	0.2	1.0	1.0
Mirex	MIREX	0.5	1.0	1.0
Fraction #1 & #2 Analytes †, ‡				
p,p'-DDE	PPDDE	1.0	1.0	0.5
p,p'-DDT	PPDDT	1.0	4.0	2.0
p,p'-DDMU	PPDDMU	2.0	5.0	5.0
trans-Nonachlor	TNONA	0.5	1.0	1.0
Fraction #2 Analytes ‡				
cis-Chlordane	CCHLOR	0.5	1.0	1.0
trans-Chlordane	TCHLOR	0.5	1.0	1.0
Chlorpyrifos	CLPYR	1.0	4.0	4.0
Dacthal	DACTH	0.2	2.0	2.0
o,p'-DDD	OPDDD	1.0	5.0	5.0
p,p'-DDD	PPDDD	0.4	3.0	3.0
p,p'-DDMS	PPDDMS	3.0	20	20
p,p'-Dichlorobenzophenone	DICLB	3.0	25	25
Methoxychlor	METHOXY	1.5	15	15
Dieldrin	DIELDRIN	0.5	1.0	1.0
Endosulfan I	ENDO_I	0.5	1.0	1.0
Endosulfan II	ENDO_II	1.0	3.0	3.0
Endosulfan sulfate	ESO4	2.0	5.0	5.0
Endrin	ENDRIN	2.0	6.0	6.0
Ethion	ETHION	2.0	NA	NA
alpha-HCH	HCHA	0.2	1.0	1.0
beta-HCH	HCHB	1.0	3.0	3.0
gamma-HCH	HCHG	0.2	0.8	1.0
delta-HCH	HCHD	0.5	2.0	2.0
Heptachlor Epoxide	HE	0.5	1.0	1.0
cis-Nonachlor	CNONA	0.5	1.0	1.0
Oxadiazon	OXAD	6	NA	NA
Oxychlordane	OCDAN	0.5	0.2	1.0

† The quantitation surrogate is PCB 103.

‡ The quantitation surrogate is d8-p,p'-DDD

Table 4: Dry Weight Detection Limits of NIST PCB Congeners.

Analytes[†]	*Database Abbreviation*	*MDL, ng/g dry Sediment*	*MDL, ng/g dry Tissue*	*MDL, ng/L Water*
2,4'-dichlorobiphenyl	PCB8	0.5	1.0	1.0
2,2',5-trichlorobiphenyl	PCB18	0.5	1.0	1.0
2,4,4'-trichlorobiphenyl	PCB28	0.5	1.0	1.0
2,2',3,5'-tetrachlorobiphenyl	PCB44	0.5	1.0	1.0
2,2',5,5'-tetrachlorobiphenyl	PCB52	0.5	1.0	1.0
2,3',4,4'-tetrachlorobiphenyl	PCB66	0.5	1.0	1.0
2,2',3,4,5'-pentachlorobiphenyl	PCB87	0.5	1.0	1.0
2,2',4,5,5'-pentachlorobiphenyl	PCB101	0.5	1.0	1.0
2,3,3',4,4'-pentachlorobiphenyl	PCB105	0.5	1.0	1.0
2,3',4,4',5-pentachlorobiphenyl	PCB118	0.5	1.0	1.0
2,2',3,3',4,4'-hexachlorobiphenyl	PCB128	0.5	1.0	1.0
2,2',3,4,4',5'-hexachlorobiphenyl	PCB138	0.5	1.0	1.0
2,2',4,4',5,5'-hexachlorobiphenyl	PCB153	0.5	1.0	1.0
2,2',3,3',4,4',5-heptachlorobiphenyl	PCB170	0.5	1.0	1.0
2,2',3,4,4',5,5'-heptachlorobiphenyl	PCB180	0.5	1.0	1.0
2,2',3,4',5,5',6-heptachlorobiphenyl	PCB187	0.5	1.0	1.0
2,2',3,3',4,4',5,6-octachlorobiphenyl	PCB195	0.5	1.0	1.0
2,2',3,3',4,4',5,5',6-nonachlorobiphenyl	PCB206	0.5	1.0	1.0
2,2',3,3',4,4',5,5',6,6'-decachlorobiphenyl	PCB209	0.5	1.0	1.0

[†] PCB 103 is the surrogate used for PCBs with 1 - 6 chlorines per molecule. PCB 207 is used for all others.

Table 5: Additional PCB Congeners with Dry Weight Minimum Detection Limits.

Analytes †	*Database Abbreviation*	*MDL, ng/g dry Sediment*	*MDL, ng/g dry Tissue*	*MDL, ng/L Water*
2,3-dichlorobiphenyl	PCB5	0.5	1.0	1.0
4,4'-dichlorobiphenyl	PCB15	0.5	1.0	1.0
2,3',6-trichlorobiphenyl	PCB27	0.5	1.0	1.0
2,4,5-trichlorobiphenyl	PCB29	0.5	1.0	1.0
2,4',4-trichlorobiphenyl	PCB31	0.5	1.0	1.0
2,2,'4,5'-tetrachlorobiphenyl	PCB49	0.5	1.0	1.0
2,3',4',5-tetrachlorobiphenyl	PCB70	0.5	1.0	1.0
2,4,4',5-tetrachlorobiphenyl	PCB74	0.5	1.0	1.0
2,2',3,5',6-pentachlorobiphenyl	PCB95	0.5	1.0	1.0
2,2',3',4,5-pentachlorobiphenyl	PCB97	0.5	1.0	1.0
2,2',4,4',5-pentachlorobiphenyl	PCB99	0.5	1.0	1.0
2,3,3',4',6-pentachlorobiphenyl	PCB110	0.5	1.0	1.0
2,2',3,3',4,6'-hexachlorobiphenyl	PCB132	0.5	1.0	1.0
2,2',3,4,4',5-hexachlorobiphenyl	PCB137	0.5	1.0	1.0
2,2',3,4',5',6-hexachlorobiphenyl	PCB149	0.5	1.0	1.0
2,2',3,5,5',6-hexachlorobiphenyl	PCB151	0.5	1.0	1.0
2,3,3',4,4',5-hexachlorobiphenyl	PCB156	0.5	1.0	1.0
2,3,3',4,4',5'-hexachlorobiphenyl	PCB157	0.5	1.0	1.0
2,3,3',4,4',6-hexachlorobiphenyl	PCB158	0.5	1.0	1.0
2,2',3,3',4,5,6'-heptachlorobiphenyl	PCB174	0.5	1.0	1.0
2,2',3,3',4',5,6-heptachlorobiphenyl	PCB177	0.5	1.0	1.0
2,2',3,4,4',5',6-heptachlorobiphenyl	PCB183	0.5	1.0	1.0
2,3,3',4,4',5,5'-heptachlorobiphenyl	PCB189	0.5	1.0	1.0
2,2',3,3',4,4',5,5'-octachlorobiphenyl	PCB194	0.5	1.0	1.0
2,2',3,3',4,5',6,6'-octachlorobiphenyl	PCB201	0.5	1.0	1.0
2,2',3,4,4',5,5',6-octachlorobiphenyl	PCB203	0.5	1.0	1.0

† PCB 103 is the surrogate used for PCBs with 1 - 6 chlorines per molecule. PCB 207 is used for all others.

Table 6: Dry Weight Minimum Detection Limits of Chlorinated Technical Grade Mixtures. Aroclors 1248, 1254, and 1260 were calculated from measured congener-specific results as described in Newman et al. (in press).

Analyte	*Database Abbreviation*	*MDL, ng/g dry Sediment*	*MDL, ng/g dry Tissue*	*MDL, ng/L Water*
Toxaphene‡	TOXAPH	50	100	100
Polychlorinated Biphenyl Aroclor 1248	ARO1248	5	100	100
Polychlorinated Biphenyl Aroclor 1254	ARO1254	5	50	50
Polychlorinated Biphenyl Aroclor 1260	ARO1260	5	50	50
Polychlorinated Terphenyl Aroclor 5460†	ARO5460	10	100	100

† The quantitation surrogate is PCB 207.

‡ The quantitation surrogate is d8-p,p'-DDD

Table 7: Dry Weight Minimum Detection Limits of Polyaromatic Hydrocarbons in Tissue

Analytes[†]	Database Abbreviation	MDL, ng/g dry Sediment	MDL, ng/g dry Tissue	MDL, ng/L Water
Naphthalene	NPH	5	10	30
2-Methylnaphthalene	MNP2	5	10	30
1-Methylnaphthalene	MNP1	5	10	30
Biphenyl	BPH	5	10	30
2,6-Dimethylnaphthalene	DMN	5	10	30
Acenaphthylene	ACY	5	10	30
Acenaphthene	ACE	5	10	30
2,3,5-Trimethylnaphthalene	TMN	5	10	30
Fluorene	FLU	5	10	30
Dibenzothiophene	DBT	5	10	30
Phenanthrene	PHN	5	10	30
Anthracene	ANT	5	10	30
1-Methylphenanthrene	MPH1	5	10	30
Fluoranthrene	FLA	5	10	30
Pyrene	PYR	5	10	30
Benz[a]anthracene	BAA	5	10	30
Chrysene	CHR	5	10	30
Tryphenylene	TRY	5	10	30
Benzo[b]fluoranthrene	BBF	5	10	30
Benzo[k]fluoranthrene	BKF	5	10	30
Benzo[e]pyrene	BEP	5	10	30
Benzo[a]pyrene	BAP	5	10	30
Perylene	PER	5	10	30
Indeno[1,2,3-cd]pyrene	IND	5	15	45
Dibenz[a,h]anthracene	DBA	5	15	45
Benzo[ghi]perylene	BGP	5	15	45
Coronene	COR	5	15	45

[†] See QA report for surrogate assignments.

Table 8: Dry Weight Minimum Detection Limits of Organometalic Compounds

Analytes[†]	*Database Abbreviation*	*MDL, ng/g dry Sediment*	*MDL, ng/g dry Tissue*	*MDL, ng/L Water*
Tributyltin	TBT	13	20	1

Total Organic Carbon Analysis of Sediments

Summary of Methods

Samples were received in the frozen state and allowed to thaw at room temperature. Source samples were gently stirred and sub-samples were removed with a stainless steel spatula and placed in labeled 20 ml polyethylene scintillation vials. Approximately 5 grams equivalent dry weight of the wet sample was sub-sampled.

Sub-samples were treated with two, 5 ml additions of 0.5 N reagent grade HCl to remove inorganic carbon (CO^{-3}), agitated, and centrifuged to a clear supernate. Some samples were retreated with HCl to remove residual inorganic carbon. The evolution of gas during HCl treatment indicates the direct presence of inorganic carbon (CO^{-3}). After HCl treatment and decanting, samples were washed with approximately 15 ml of deionized-distilled water, agitated, centrifuged to a clear supernate, and decanted. Two sample washings were required to remove weight determination and analysis interferences.

Prepared samples were placed in a 60° C convection oven and allowed to come to complete dryness (approx. 48 hrs.). Visual inspection of the dried sample before homogenization was used to ensure complete removal of carbonate containing materials, (shell fragments). Two 61 mm (1/4") stainless steel solid balls were added to the dried sample, capped and agitated in a commercially available ball mill for three minutes to homogenize the dried sample.

A modification of the high temperature combustion method, utilizing a Weatstone bridge current differential was used in a commercially available instrument, (Control Equipment Co., 440 Elemental Analyzer) to determine carbon and nitrogen concentrations. The manufactures suggested procedures were followed. The methods are comparable to the validation study of USEPA method MARPCPN I. Two to three aliquots of 5-10 mg of dried prepared sub-sample were used to determine carbon and nitrogen weight percent values. Calibration of the instrument was with known standards using Acetanilide or L-Cystine. Detection limits are 0.2 ug/mg, carbon and 0.01 ug/mg nitrogen dry weight.

The above methods and protocols are modifications of several published papers, reference procedures and analytical experimentation experience (Franson, 1981; Froelich, 1980; Hedges and Stern, 1983; MARPCPN I, 1992).

Quality Control/Quality Assurance for TOC Measurement

Quality control was tested by the analysis of National Research Council of Canada Marine Sediment Reference Material, BCSS-1 at the beginning and end of each sample analysis set (20-30 individual machine analyses). All analyzed values were within suggested criteria of $\pm$ 0.09% carbon (2.19% Average). Nitrogen was not reported on the standard data report, but was accepted at $\pm$ 0.008% nitrogen (0.195% Average) from the EPA study. Quality assurance was monitored by re-calibration of the instrument every twenty samples and by the analysis of a standard as a unknown and comparing known theoretical percentages with resultant analyzed percentages. Acceptable limits of standard unknowns were less than $\pm$ 2%. Duplicate or triplicate sample analysis variance (standard

deviation/mean) greater than 7% is not accepted. Samples were re-homogenized and re-analyzed until the variance between individual runs fell below the acceptable limit of 7.0%.

Grain Size Analysis of Sediments

Summary of Methods

The procedure used combined wet and dry sieve techniques to determine particle size of sediment samples. Methods follow those of Folk (1974).

Sample Splitting and Preparation

Samples were thawed and thoroughly homogenized by stirring with a spatula. Spatulas were rinsed of all adhering sediment between samples. Size of the subsample for analysis was determined by the sand/silt ratio of the sample. During splitting, the sand/silt ratio was estimated and an appropriate sample weight was calculated. Subsamples were placed in clean, pre-weighed beakers. Debris was removed and any adhering sediment was washed into the beaker.

Wet Sieve Analysis (Separation of Coarse and Fine Fraction)

Beakers were placed in a drying oven and sediments were dried at less than 55°C until completely dry (approximately three days). Beakers were removed from drying oven and allowed to equilibrate to room temperature for a least a half-hour. Each beaker and its contents were weighed to the nearest .01 g. This weight minus the empty beaker weight was the total sample weight. Sediments in beakers were disaggregated using 100 ml of a dispersant solution in water (such as 50g Calgon/L water) and the sample was stirred until completely mixed and all lumps disappear. The amount and concentration of dispersant used was recorded on the data sheet for each sample. Sample beakers were placed in an ultrasonic cleaner for 15 minutes for disaggregation. Sediment dispersant slurry was poured into a 63 μm (ASTM #230, 4 phi) stainless steel or brass sieve in a large glass funnel suspended over a 1L hydrometer cylinder by a ring stand. All fine sediments were washed through the sieve with water. Fine sediments were captured in a 1L hydrometer cylinder. Coarse sediments remaining in sieve were collected and returned to the original sample beaker for quantification.

Dry Sieve Analysis (Coarse Fraction)

The coarse fraction was placed into a preweighed beaker, dried at 55-65°C, allowed to acclimate, and then weighed to 0.01 g. This weight, minus the empty beaker weight, was the coarse fraction weight. The coarse fraction was poured into the top sieve of a stack of ASTM sieves having the following sizes: No. 10 (2.0 mm), 18 (1.0 mm), 45 (0.354 mm), 60 (0.25 mm), 80 (0.177 mm), 120 (0.125 mm), and 170 (0.088 mm). The stack was placed on a mechanical shaker and shaken at medium intensity for 15 minutes. After shaking, each sieve was inverted onto a large piece of paper and tapped 5 times to free stuck particles. The sieve fractions were added cumulatively to a weighing dish, and the cumulative weight after each addition determined to 0.01g. The sample was returned to its original beaker, and saved until sample computations were completed and checked for errors.

Analytical Procedures

Fractional weights and percentages for various particle size fractions were calculated. If only wet sieve analysis was used, weight of fine fraction was computed by subtracting coarse fraction from

total sample weight, and percent fine composition was calculated using fine fraction and total sample weights. If dry sieve was employed as well, fractional weights and percentages for the sieve were calculated using custom software on a Macintosh computer. Calibration factors were stored in the computer.

Hydrometer Analysis (Fine Fraction)
Hydrometers used for the analysis were precalibrated using the techniques of Lewis (1984). A reference cylinder was filled with water and 100 ml of dispersant solution. Prior to the analysis, a hydrometer reading was taken for Cc, the composite correction for temperature, dispersing agent, and the meniscus.

For each of the sample cylinders, the volume was raised to 1000 ml using tap water. The hydrometer number was recorded, the temperature was noted, and the sample added and stirred for 1 minute.

Hydrometer readings were taken at 1 minute, 3 minutes, 10 minutes, 30 minutes, 90 minutes, 4.5 hours and 24 hours. If the water temperature had changed by greater than 2°C then hydrometer corrections were remeasured. The colloidal weight was determined by subtracting the other fractions from the total weight.

Benthic Community Analysis

Summary of Methods
Each catalogued sample was processed individually in the laboratory to obtain an accurate assessment of species diversity and abundance. All macroinvertebrates were sorted from residues under a dissecting microscope, identified to lowest possible taxon, and counted. Laboratory processing of benthic cores consists of both rough and fine sorting. Initial sorting separates animals into large taxonomic groups such as polychaetes, crustaceans, mollusks and other (*e.g.*, phoronids). Bound laboratory logbooks were maintained and used to record number of samples processed by each technician, as well as results of any sample resorts, if necessary. Sorters were required to sign and date a Milestone Progress Checksheet for each replicate sample processed. Specimens of similar taxonomic groups were placed in vials and labeled internally and externally with project, date collected, site/station information, and IDORG. Samples were selected for benthic community analysis by SWRCB staff based on results from toxicity tests.

In-house senior taxonomists and outside specialists processed and verified the accuracy of species identification and enumeration. An archived voucher specimen collection was established at this time.

Relative Benthic Index
Benthic samples were sieved, sorted and the number of individuals of each species in each replicate core were identified. A number of summary statistics were calculated for each station, including summaries of total fauna, number of species, and the 4 major phyla (Polychaetes, Crustaceans, Mollusks, and Echinoderms).

The Relative Benthic Index (RBI) used in this study utilizes the above summarized fauna information in a refined version of the benthic index presented in the San Diego BPTCP report (Fairey et al. 1996). It is based on simple, realistic natural history concerning responses of marine benthic communities to anthropogenic and natural disturbances. The community patterns used in the index include number of species (all taxa, only mollusks, and only crustaceans); and the number of individuals of crustaceans, the number of individuals of selected species that are indicators of relatively disturbed benthic habitats, and the number of individuals of selected species that are indicators of relatively undisturbed benthic habitats. The RBI is developed for particular areas by selecting different indicator species. It does not require the presence of uncontaminated reference stations, and does not refer to data beyond that collected in each study. Often the evaluation of community degradation depends on comparisons to uncontaminated reference sites which are difficult to locate and vary for reasons that are unknown and unrelated to contamination.

Number of Species

The number of species often decreases with severe disturbances (Oliver et al. 1977, 1980, Lenihan and Oliver 1995) and is the best indicator of biodiversity, particularly when species are sampled in relation to habitat area (Hurlbert 1971, Jumars 1975, 1976, Abel and Walters 1979). Therefore, the first community parameter in the RBI is the total number of species found in a standard sample of habitat area. Among the more numerous large taxonomic groups, crustaceans are generally more sensitive to environmental contaminants and other anthropogenic disturbances than most other components of the infauna, particularly polychaetes (Pearson and Rosenberg 1978, Reish et al. 1980, Thistle 1981, Swartz et al. 1986, Stull et al. 1986, Oliver et al. 1977, Lenihan and Oliver 1995, Lenihan et al. 1995). Speciose and numerically abundant crustacean faunas on the Pacific coast of the United States are generally only found in uncontaminated environments (Barnard 1963), making the number of crustacean species an important indicator of overall environmental health. To a lesser degree, the number of mollusk species also increase with decreasing environmental stress (Stull et al. 1986, Swartz et al. 1986, Oliver et al. 1977), and are thus also included in the RBI. Polychaetes, crustaceans, and mollusks are the three dominate groups of benthic macro-invertebrates from many nearshore communities (Oliver et al. 1980), but unlike the crustaceans and mollusks, many of the most opportunistic or weedy species are polychaete (Grassle and Grassle 1974, McCall 1977, Oliver et al. 1977, Pearson and Rosenberg 1978, Reish et al. 1980, Sanders et al. 1980, Santos and Simon 1980, Thistle 1981, Rhoads et al. 1982, Lenihan and Oliver 1995). As a result, the number of polychaete species was not used in the RBI, because they do not indicate as clearly either a relatively disturbed habitat or a relatively undisturbed habitat.

Number of Individuals

An increase in the number of crustacean individuals is also indicative of relatively healthy environments (Stull et al. 1986, Swartz et al. 1986, Oliver et al. 1977, Lenihan and Oliver 1995), although sometimes one or two crustacean species can be abundant in disturbed habitats (Vetter 1995, Okey 1997), but less so than for other major taxonomic groups, particularly polychaete worms (Pearson and Rosenberg 1978, Grassle and Grassle 1974, Oliver et al. 1977). Therefore, the number of individuals of crustaceans is also used in the RBI, but not the number of

individuals in any other major taxonomic group.

Indicator Species

Even more than the number of species or the number of crustacean individuals, the population sizes of selected indicator species are strongly associated with benthic habitats that are relatively disturbed or undisturbed (Grassle and Grassle 1974, Oliver et al. 1977, Davis and Spies 1980, Westin 1990, Lenihan and Oliver 1995, Okey 1997). Therefore, five species were used in the RBI as indicators of either highly disturbed or undisturbed benthic communities and habitats. The number and identity of indicator species can change from one regional study site to another. Selection of indicator species was based on known responses to anthropogenic and other disturbances (Grassle and Grassle 1974, McCall 1977, Oliver et al. 1977, Pearson and Rosenberg 1978, Davis and Spies 1980, Sanders et al. 1980, Santos and Simon 1980, Thistle 1981, Lenihan and Oliver 1995, Okey 1997) and related natural history such as life history traits (Grassle and Grassle 1974, Oliver et al. 1977, Rhoads et al. 1978, Rhoads and Boyer 1982, Lenihan and Oliver 1995) and abundance patterns along environmental gradients and among the study stations (Oliver et al. 1980, Stull et al. 1986, Swartz et al. 1986, Weston 1990). The 2 negative indicator species are highly opportunistic annelids which thrive in disturbed, polluted, or marginal environments, and are generally not found in less disturbed communities. The 3 positive indicator species are generally not found in polluted habitats and are characteristic of regions where anthropogenic and other severe disturbances do not play major roles in structuring communities. Each indicator species is discussed below:

Negative Indicator Species

Capitella capitata

The *Capitella* species complex is a cosmopolitan group which lives in a wide range of conditions: fouled or low oxygen, high organic matter and fine sediments. They are abundant around outfalls discharging biological wastes, and have a rapid (1 to 2 month) life cycle. *Capitella* are capable of surviving for days with little or no oxygen, and are often considered the best example of a "weedy", opportunistic species (Grassle and Grassle 1974, Grassle and Grassle 1976, Oliver et al. 1977, McCall 1977, Pearson and Rosenberg 1978, Lenihan and Oliver 1995, Okey 1995 and many others).

Oligochaetes

Oligochaetes are a poorly known group typically found in peripheral/disturbed habitats such as under decaying algae on beaches, and in fouled or low oxygen muds of back bays, estuaries, and harbors (Brinkhurst and Simmons 1968, Pearson and Rosenberg 1978, Brinkhurst and Cook 1980). They often occur in large masses with nearly no other macrofauna. In SF Bay they may comprise 100% of the fauna where there is gross pollution (i.e. large amounts of organic material from sewage). If oxygen levels are sufficient, and there is little toxic waste and high bacterial levels, oligochaete densities become extremely high (Smith and Carlton, 1975; Brinkhurst and Simmons, 1968). They are well known indicators of relatively degraded freshwater ecosystems (Brinkhurst and Simmons 1968, Pearson and Rosenberg 1978, Brinkhurst and Cook 1980).

Positive Indicator Species

Ampelisca

Ampelisca filter-feed from vertical tubes which they build at the surface of clean, fine sediments. Tremendous densities of *Ampelisca* can form a dense carpet of tubes which changes the physical structure of the sedimentary regime. The carpet also enhances habitat values and supports a very diverse fauna (Mills 1967, Oliver et al. 1983, 1984, Oliver and Slattery 1985). Although *Ampelisca* can colonize open sediment patches (Mills 1967), they do not colonize disturbed sites nearly as rapidly as the more motile and non-tube dwelling amphipod groups (Oliver and Slattery 1985b, Klaus et al. 1990).

Macoma

The clams *Macoma* and *Tellina*, both in the Tellinidae, are small and live shallowly under the sediment surface. *Macoma* generally favors finer sediment, including bays, than *Tellina*. Some *Macomas* filter feed, others deposit feed by vacuuming sediment surface with their incurrent siphon (Reid and Reid 1969). They are not known to be early colonists in disturbed sedimentary habitats (Oliver et al. 1977).

Tellina

Tellina live in clean, well-oxygenated sands of shallow water (Oliver et al. 1980). Species in Southern California attain great enough densities to be a major component of the shallow water, benthic infaunal community (Barnard 1963). They are not known to be early colonists in disturbed sedimentary habitats (Oliver et al. 1977).

Calculation of Relative Benthic Index

Previous versions of the RBI have used individual impact thresholds for determination of degree of negative impact to Total Fauna and Number of Crustacean Species (Fairey et al. 1996). While these thresholds have been useful, the necessarily arbitrary nature of the selection process introduced potential artifacts for stations whose values for Total Fauna, Total Mollusks and Total Crustacea approached the threshold value. To address this problem, calculation of the RBI was revised to be based on percentages of the total range. The final threshold value for determination of impacted versus non-impacted sites was based on the overall RBI and selected using best professional judgment. Justification for this critical threshold value of the RBI is discussed below.

For total fauna, number of mollusk species and number of crustacean species, the maximum and minimum values in these parameters over all the stations were determined. For each station, the total number of species, total mollusk species, and total number of crustacean species were then converted to the percentage of the total range for these parameters. The number of crustacean individuals at each station is similarly converted to a percentage of the total range, and is added to the total fauna, mollusk, and crustacean species numbers. The community numbers thus represent four sixths of the Relative Benthic Index for each station.

For the positive and negative indicator indices, the final index was weighted towards presence and absence of key indicator species, with abundance of each species given additional incremental weight. Accordingly, the abundance of each indicator species was transformed using

a double square-root transformation to compress the range of values. For each species, the transformed abundance was converted to a percentage of the total range. The transformed values of the negative indicator species were summed and subtracted from the sum of the values for the positive indicator species.

The overall RBI was calculated by summing the values of the Total Fauna, Total Mollusks, Crustacean Species, and Indicator Species, and standardizing it to the total range. This resulted in a range in values from 0.00 (Most Impacted) to 1.00 (Least Impacted).

Use of Relative Benthic Index

It is not possible to compare directly RBI values between different regions. The high and low ranges of values vary based on the extreme values within each data set. In addition, different indicator species are often used between regions. What the RBI does provide is the relative "health" of each of the stations in a given data set compared to the other stations in the same data set.

The RBI does not indicate causality. While a low RBI value could be the result of chemical toxicity, it also could be the result of other types of anthropogenic disturbance, such as dredging, or could result from a variety of natural disturbances, such as freshwater runoff, temperature stratification, or storm impacts.

It is not possible to test the RBI to determine significance levels or confidence levels, or to statistically determine what ranking indicates significant impact. However, since a degree of arbitrarity is incorporated into all determinations of significance, whether statistical or intuitive, this should not be considered a significant drawback. For this study, the threshold for significantly impacted benthic community structure was set at a RBI less than or equal to 0.30. While this threshold is necessarily somewhat arbitrary, it is considered suitable based on the best professional judgment of the benthic ecologists who performed the analysis. Several factors were considered in deriving this threshold: the stations below the threshold have few overall species, few crustacean species, presence of negative indicator species, and absence of positive indicator species. These stations would be considered to be significantly degraded by the vast majority of naturalists familiar with San Francisco Bay. The RBI can be used in combination with chemistry and toxicity test data to provide a "weight-of-evidence" for determination of the most impacted stations.

Bioaccumulation Tests with the Clam Macoma balthica

The 28-day bioaccumulation tests with *Macoma balthica* were conducted according to EPA/Army Corps of Engineers Inland Testing Manual (EPA/ACOE, 1994). Clams were obtained from Brezina and Associates (Dillon Beach, CA). Clams arrived via overnight courier on day 0 of the test. Test containers consisted of 5L polyethylene trays with 2.5L of sediment. Sediment was loaded into test containers, and allowed to equilibrate for 24 hours before clams were added. Fifteen clams were placed in each of 3 replicate containers and flow-through seawater was started at a rate of 120 mL per minute. After 28 days, sediment was screened and discarded. Surviving clams were placed in clean, flow-through seawater to depurate for 24 hours. After depuration,

clams were blotted dry, weighed, and frozen for tissue analysis at minus 12°C.

A negative control consisting of either Yaquina Bay amphipod home sediment from NWAS or *Macoma* collection site sediment was used. Three replicates of clams were depurated for 24 hours at the initiation of the test to obtain baseline tissue concentrations.

Statistical Analyses

Toxicity Data

The statistical significance of toxicity test results was evaluated using three complimentary methods. Two of the methods were based on comparison to laboratory controls, while the third was based on comparison to reference sites.

Comparisons to Laboratory Controls

Samples were defined as significantly more toxic than laboratory controls if the following two criteria were met: 1) a separate-variance t-test determined there was a significant difference ($p < 0.05$) in mean toxicity test organism response (e.g., percent survival) between the sample and the laboratory control, and 2) mean organism response in the toxicity test was lower than a certain percentage of the control value, as determined using the 90th percentile Minimum Significant Difference (MSD).

Statistical significance in t-tests is determined by dividing an expression of the difference between sample and control by an expression of the variance among replicates. We used a "separate variance" t-test that adjusted the degrees of freedom to account for variance heterogeneity among samples. If the difference between sample and control was large relative to the variance among replicates, then the difference was considered significant. In many cases, however, low between-replicate variance will cause a comparison to be considered significant, even though the magnitude of the difference can be small. The magnitude of difference that can be identified as significant is termed the Minimum Significant Difference (MSD), which is dependent on the selected alpha level, the level of between-replicate variation, and the number of replicates specific to the experiment. With the number of replicates and alpha level held constant, the MSD varies with the degree of between-replicate variation. The "detectable difference" inherent to the toxicity test protocol can be determined by identifying the magnitude of difference that can be detected by the protocol 90% of the time (Schimmel et al., 1994; Thursby and Schlekat, 1993). This is equivalent to setting the level of statistical power at 0.90 for these comparisons. This is accomplished by determining the MSD for each t-test conducted, ranking them in ascending order, and identifying the 90th percentile MSD (the MSD that is larger than or equal to 90% of the MSD values generated).

Thursby et al. (1997) identify a value of 80% of the control as the detectable difference for the *Ampelisca* test, and similar values for other species have been derived from BPTCP test data. Current BPTCP detectable difference (90th percentile MSD) values are listed in Table 9. Samples with toxicity test results lower than the values given, as a percentage of control response, would be considered toxic if the result was also significantly different from the control in the individual t-test.

Table 9. Ninetieth percentile MSD values and threshold percentage of control values used in determining statistically significant sample toxicity. MPSL indicates values derived by the Marine Pollution Studies Laboratory, using the entire BPTCP data base. *Protocols for which insufficient BPTCP data were available were assigned the 80% value derived for *Ampelisca* by Thursby et al. (1997).

Protocol	MSD	% of control	N	Reference
Ampelisca solid-phase	20	80		Thursby, 1997
Ceriodaphnia (96-h) porewater	20	80		Thursby, 1997*
Ceriodaphnia (96-h) SWI	20	80		Thursby, 1997*
Eohaustorius solid-phase	25	75	385	MPSL
Hyalella solid-phase	20	80		Thursby, 1997*
Abalone water (5 reps)	10	90	131	MPSL
Abalone water (3 reps)	36	64	336	MPSL
Abalone water (all reps)	32	68	467	MPSL
Mytilus porewater	20	80	223	MPSL
Neanthes Surv. solid-phase	36	64	335	MPSL
Neanthes Wt. solid-phase	56	44	335	MPSL
Rhepoxynius solid-phase	23	77	720	MPSL
Urchin Dev. porewater (5 reps)	22	78	309	MPSL
Urchin Dev. porewater (3 reps)	45	55	630	MPSL
Urchin Dev. porewater (all)	40	60	939	MPSL
Urchin Dev. SWI	41	59	109	MPSL
Urchin Fertilization	12	88	79	MPSL

Reference Envelope Statistical Method

The "reference envelope" approach was developed to provide an appropriate statistical method for determining whether conditions at test sites were significantly worse than those in the surrounding area. This objective is different from that of determining absolute sample toxicity, as described above. Rather than comparing results of test samples with laboratory controls using laboratory replicate variance as the statistical test variance component, the reference envelope method establishes tolerance limits based on test results from reference site samples. Tolerance limits are calculated to identify samples significantly more toxic than a chosen proportion of the reference site distribution, and statistical significance is determined using variation among reference site results. In this way, the method considers all relevant sources of variation that could affect comparisons between sites, such as variation in time and space, the interaction of time and space components, and variation between replicates (the error term). If natural factors such as grain size vary among reference sites or between surveys, then the effects of these factors are accounted for in the analysis. Any additional variation (i.e., increased toxicity) is assumed to be the result of increased pollution at test sites.

Tolerance limits were calculated using reference site data collected during a study specifically designed to evaluate reference sites in San Francisco Bay (Hunt et al., 1998), and detailed

information on reference sites and the reference envelope statistical method were presented there. In this study, samples with survival or normal development less than the respective tolerance limits were considered toxic. Tolerance limits used in the present study were based on an alpha value of 0.05 and a "p" value of 10. Samples would thus be considered toxic if there were at least a 95% probability that the sample was as toxic or more toxic than would be expected of the worst 10% of reference site samples. Tolerance limits for each toxicity test protocol used in this study, as a percentage of test control values, are given in Table 10.

Table 10. Reference Envelope Tolerance Limits.

Protocol	Tolerance Limit as % of the control
Amphipod (*Eohaustorius*) Survival	69.5 %
Sea Urchin Larval Development in Porewater	94.3 %
Sea Urchin Larval Development at SWI	86.7 %

The high values for the sea urchin protocols indicate that reference site samples had very low toxicity to sea urchin larvae, with little variation among samples. These high sea urchin test tolerance limits do not necessarily indicate that these differences from the control were biologically significant. For the sea urchin tests, we have deferred to the lower standard defined by the detectable difference values given in Table 9.

Chemistry Data

Comparisons with Sediment Quality Guideline Values

Bioavailability is the key to understanding the relationship between sediment chemistry and biological impacts. However, it was not possible to use TIEs, bioaccumulation analyses, or other specialized methods to evaluate bioavailability on the large number of samples evaluated in BPTCP studies to date. In order to assess large numbers of samples for their potential to impact biological resources, we compared sediment chemical concentrations to published guideline values derived from studies of approximately one thousand samples collected nationwide. These studies have used empirical observation of large data sets containing matching chemistry and biology data to provide guidance for evaluating the probability that measured contaminant concentrations might contribute to observed biological effects (MacDonald, 1994; Long et al., 1995). While the reported guideline values were derived from sediments containing mixtures of chemicals, they were calculated individually for each chemical. Their application may be confounded in sediments where biological responses are affected by synergistic or antagonistic interactions among multiple compounds, by unmeasured or unidentified compounds, or by unmeasured physical factors.

The National Status and Trends Program has evaluated chemical and toxicological evidence from a number or laboratory, field, and modeling studies to establish three ranges of chemical concentrations that are either rarely, sometimes, or usually associated with biological effects. Evaluation of available data (Long et al., 1995) has resulted in the identification of three

concentration ranges for selected chemical compounds:

1) Minimal Effects Range: The range in concentrations over which toxic effects are rarely observed.

2) Possible Effects Range: The range in concentrations over which toxic effects are occasionally observed.

3) Probable Effects Range: The range in concentrations over which toxic effects are frequently or always observed.

Two different methods were used to determine these chemical ranges. One method developed by NOAA (Long et al., 1995) used chemical data which were associated with a toxic response. These data were used to determine the lowest 10th percentile of ranked data where chemical concentration was associated with an effect (Effects Range - Low, or ERL). Chemical concentrations below the ERL are expected to rarely affect organisms. The Effects Range-Median (ERM) reflects the 50th percentile of ranked data and represents the level above which effects are expected to occur. Effects are occasionally expected to occur when chemical concentrations fall between the ERL and ERM.

The screening concentrations described by MacDonald (1996) also identify three ranges of chemical concentrations associated with toxic biological responses, but use an alternate method. The ranges are differentiated by the PEL (Probable Effects Level) and TEL (Threshold Effects Level). TELs were derived by taking the geometric mean of the 50th percentile of the "No Effects" data and the 15th percentile of the "Effects" data. The PEL values were derived by taking the geometric mean of the 85th percentile of the "No Effects" data and the 50th percentile of the "Effects" data. The ERL, ERM, TEL, and PEL values are provided in Table 11.

Although different data sets and percentiles were used in these two approaches to derive chemical screening concentrations, they are in close agreement, usually within a factor of 2. While neither of these methods is advocated over the other in this report, we have presented only ERM comparisons to simplify the many presentations of the data. Long, Field, and MacDonald (1998) found that the predictive ability of ERMs was slightly greater than that of PELs in a recent evaluation of additional sediment data.

It should be noted that the degree of confidence that MacDonald (1996) and Long et al. (1995) had in their respective numerical guidelines varied considerably among the different chemicals. For example, neither had great confidence in the values for nickel, mercury, DDTs, dieldrin, and endrin. DDT compounds were among those exceeding the PEL and ERM values most often at the 43 stations sampled in this study. Due to the pervasive presence of DDT compounds and the uncertainty of the DDT ERM value, we have used an alternative DDT guideline value in this report. That value is 100 μg DDT per gram organic carbon, derived by Swartz et al. (1994) from intensive studies in Lauritzen Channel, Richmond Harbor, San Francisco Bay.

Table 11. Sediment Quality Guideline values developed by NOAA and the State of Florida.

	State of Florida (1)		NOAA (2,3)	
SUBSTANCE	TEL	PEL	ERL	ERM
Total PCB (ug/kg- dry weight)	21.550	188.79	22.70	180.0
PAH (ug/kg- dry weight)				
Acenaphthene	6.710	88.90	16.00	500.0
Acenaphthylene	5.870	127.89	44.00	640.0
Anthracene	46.850	245.00	85.30	1100.0
Fluorene	21.170	144.35	19.00	540.0
2-methylnaphthalene	20.210	201.28	70.00	670.0
Naphthalene	34.570	390.64	160.00	2100.0
Phenanthrene	86.680	543.53	240.00	1500.0
Total LMW-PAHs	311.700	1442.00	552.00	3160.0
Benz(a)anthracene	74.830	692.53	261.00	1600.0
Benzo(a)pyrene	88.810	763.22	430.00	1600.0
Chrysene	107.710	845.98	384.00	2800.0
Dibenz(a,h)anthracene	6.220	134.61	63.40	260.0
Fluoranthene	112.820	1493.54	600.00	5100.0
Pyrene	152.660	1397.60	665.00	2600.0
Total HMW-PAHs	655.340	6676.14	1700.00	9600.0
Total PAHs	1684.060	16770.54	4022.00	44792.0
Pesticides (ug/kg- dry weight)				
p,p'DDE	2.070	374.17	2.20	27.0
p,p'DDT	1.190	4.77		
Total DDT	3.890	51.70	1.58	46.1
Total DDT (Swartz et al., 1994)				100µg/g OC
Lindane	0.320	0.99		
Chlordane	2.260	4.79	2.00	6.0
Dieldrin	0.715	4.30		8.0
Endrin				45.0
Metals (mg/kg- dry weight)				
Arsenic	7.240	41.60	8.20	70.0
Antimony			2.00	25.0
Cadmium	0.676	4.21	1.20	9.6
Chromium	52.300	160.40	81.00	370.0
Copper	18.700	108.20	34.00	270.0
Lead	30.240	112.18	46.70	218.0
Mercury	0.130	0.70	0.15	0.7
Nickel	15.900	42.80	20.90	51.6
Silver	0.733	1.77	1.00	3.7
Zinc	124.000	271.00	150.00	410.0

(1) D.D. MacDonald, 1994
(2) Long et al., 1995
(3) Long and Morgan, 1990

Non-Guideline Chemicals

To evaluate chemicals for which no ERM or PEL guidelines have been calculated, concentrations of specific chemicals were compared to the range of chemical concentrations in the BPTCP statewide database. This database contains concentrations of approximately 120 analytes measured in sediments collected in the majority of California bays, estuaries, lagoons and coastal areas. The following information was described for each chemical: the Method Detection Limit (MDL), the number of samples analyzed, the number of samples above the MDL, the highest value in the data set, and the 90th and 95th percentile thresholds for each chemical. In this report, chemicals for which no sediment quality guideline values have been calculated were compared to the 90th and/or 95th percentile of the statewide database (Table 12).

Reference Envelope Tolerance Limits for Selected Chemicals

Tolerance limits were calculated for a number of chemicals, based on the distribution of chemical concentrations measured at reference sites in San Francisco Bay (Smith, 1997, report to the SFBRWQCB). Reference sites and the reference envelope approach were described by Hunt et al. (1998), and the reference envelope approach was briefly described above with reference to toxicity data. The calculated chemical tolerance limits are given in Table 12. These were calculated to provide 95% certainty that measured concentrations exceeding the tolerance limit would be as high or higher than expected of the highest 15% of samples from reference sites. This reflects the "p" value of 0.85 selected by the Regional Board staff when they derived threshold values for ambient concentrations of these chemicals in their assessments of test sites (Gandesbery and Hetzel, 1998). Concentrations above the tolerance limits could therefore be assumed to be elevated relative to optimal ambient conditions in the Bay. No assumptions are made about the relationship between the tolerance limit concentrations and their potential for biological effects; they are simply descriptive of chemical concentrations found at reference sites.

These values were not used in the analysis of screening and confirmation data as described in the Results and Discussion section of this report, but two points should be mentioned regarding the values in Table 12. First, for the majority of chemicals for which S.F. Bay reference tolerance limits were derived, these limits were much lower than either the 90th percentile of the BPTCP statewide distribution or the ERM values calculated using concentrations relative to observed biological effects. Second, the nickel concentration at the 85th percentile of the SF Bay reference site distribution (the tolerance limit) was higher than the 90th percentile for all BPTCP samples statewide, many of which were collected to characterize potentially polluted sites. As mentioned in the Introduction, geologic abundance and human-enhanced transport of this element, among other factors, has apparently resulted in elevated concentrations throughout San Francisco Bay.

Table 12. Chemical comparison values: the 90th percentile of the statewide BPTCP data base, reference envelope tolerance limit (p = 0.85), and ERM. MDL is minimum detection limit. Metals are in mg/kg; organics are in µg/kg. Tolerance limits assume a grain size 40-100% fines.

Chemical Name	MDL	Samples Analyzed	Highest Value	90th % Threshold	Tolerance Limit p =.85	ERM (Long et al., 1995)
Aluminum	1	603	165,000	83,000	na	n/a
Antimony	0.1	603	52.8	3.35	na	25
Arsenic	0.1	544	1140	21.2	15.3	70
Cadmium	0.002	603	27.9	1.76	0.33	9.6
Chromium	0.02	603	860	212	112	370
Copper	0.003	603	7,800	300	68.1	270
Iron	0.1	603	336,300	55,300	na	n/a
Lead	0.03	603	2100	120	43.2	218
Manganese	0.05	603	1190	630	na	n/a
Mercury	0.03	603	9.14	0.969	0.43	0.7
Nickel	0.1	550	167	88	112	51.6
Silver	0.002	603	35.7	1.58	0.58	3.7
Selenium	0.1	544	35.7	1.09	0.64	n/a
Tin	0.02	603	92.9	9.03	na	n/a
Zinc	0.05	603	6,000	490	158	410
Aldrin	0.5	621	8.2	4.7	na	n/a
Chloropyrifos	1	444	78	28	na	n/a
Total Chlordane	3	612	246	44.57	1.1	6
Dacthal	0.2	465	25.2	7.51	na	n/a
Total DDT (*Swartz)	5.4	621	3,569	235.5	7.0	100µg/gOC*
pp-Dichlorobenzophenone	3	465	63.3	30.6	na	n/a
Dieldrin	0.5	618	62.6	11.7	0.44	8
Endosulfan I	0.5	606	19.6	13.4	na	n/a
Endosulfan II	1	606	59.8	10.4	na	n/a
Endosulfan Sulfate	2	606	163	21	na	n/a
Endrin	2	618	21.8	16.4	na	45
Ethion	2	69	36.4	36.4	na	n/a
alpha-HCH	0.2	465	292	26.1	na	n/a
beta-HCH	1	465	56.8	56.8	na	n/a
gamma-HCH (Lindane)	0.2	618	8.4	2.82	na	0.99 (PEL)
delta-HCH	0.5	465	99.4	14.4	na	n/a
Heptachlor	0.5	621	15.8	4.5	na	n/a
Heptachlor Epoxide	0.5	618	17.8	2.5	na	n/a
Hexachlorobenzene	0.2	621	59.7	3.63	0.48	n/a
Methoxychlor	1.5	606	131	55.3	na	n/a
Mirex	0.5	620	103	2.6	na	n/a
Oxadiazon	6	465	114	45.8	na	n/a
Oxychlordane	0.5	465	30.3	10.7	na	n/a
Toxaphene	50	609	3,200	3,200	na	n/a
Tributyltin	0.003	555	6.21	0.422	na	n/a
Total PCB	9	684	19,901	497	14.8	180
Low MW PAHs	60	624	92,097	2,585	434	3,160
High MW PAHs	60	628	225,740	15,727	3060	9,600
Total PAHs	60	628	227,801	17,107	3390	44,792
Total Organic Carbon	n/a	686	26.8	3	na	n/a
Mean ERM Quotient	n/a	548	4.37	1.11	na	n/a
Mean PEL Quotient	n/a	553	7.8	1.52	na	n/a

ERM and PEL Quotients

Sediment Quality Guideline quotients (SQGQ) were calculated to allow a simple comparison between observed chemical concentrations and guideline values developed for that chemical using a nationwide data base. To derive these quotients for a given sample, the concentration of each chemical was divided by its respective SQG value to get a quotient. Quotient values greater than 1 indicated that the chemical in that sample exceeded its guideline value, and was likely to be associated with biological effects, based on comparisons to the large data sets from which the guidelines were derived.

In screening samples for potential effects of chemical mixtures, the quotient values for 16 chemicals were averaged to get a mean SQG quotient. In this report, sample chemical concentrations were compared primarily to ERM values, where possible, and mean ERMQ values were generally used as summary quotients for chemical mixtures. This mean value was calculated somewhat differently from mean ERMQ values presented by Long et al. (1998), as is discussed below in the section on the use of threshold values. The 16 chemicals used to derive this mean value were: antimony, arsenic, cadmium, chromium, copper, lead, mercury, silver, zinc, total DDT (using the DDT value of Swartz, et al., 1994), total chlordane, dieldrin, endrin, total PCBs, low molecular weight (LMW) PAHs, and high molecular weight (HMW) PAHs. In cases where concentrations of these chemicals were below the analytical method detection limit (MDL), a value of one-half the MDL was used in the derivation of the mean ERMQ. The use of mean ERMQ values was designed to assist with screening samples in which multiple compounds contributed to the overall level of chemical pollution, and was intended for use in conjunction with the standard chemical-specific method discussed above. Although synergistic effects are possible with the different contaminants, this is not implied by the use of mean SQGQs.

Multivariate and Univariate Techniques for Comparison of Chemistry and Toxicity Data

While the main objective of this study was to identify stations of concern, the data were also evaluated to investigate whether certain individual chemicals were found to be associated with biological impacts. These evaluations were made using Principal Components Analysis (a multivariate technique) followed by Correlation Analysis (a univariate technique). This association of chemicals with toxicity does not in itself prove cause and effect, but it allows the suggestion of hypotheses regarding the chemical causes of biological impacts, which can later be tested with TIEs and other more extensive toxicological methods.

Principle Components Analysis

Because many chemicals tend to co-vary in sediments, Principal Components Analysis (PCA) was used to investigate relationships between chemistry, toxicity, and benthic indicators prior to conducting simple correlation analyses. The PCA was treated as exploratory in nature; therefore, data were not screened for sample size, normality, linearity, outliers or multicolinearity.

Principal components were extracted using SYSTAT statistics software (v. 7.0.1 for Windows; SPSS, 1997). The analysis was run with a correlation matrix and varimax rotation, and included any factors which accounted for greater than 10% of the total variance. A component loading

cutoff value of 0.40 was used in selecting variables for inclusion into factors, based on suggestions by Tabachnick and Fidell (1996) that a cut-off of at least 0.32 be used.

Correlation Analysis

Compounds determined by PCA to have a negative relationship with biological indicators (e.g., increasing concentration associated with decreasing survival) were selected for univariate correlation analysis. In order to examine associations between levels of these pollutants in sediments and the response observed in toxicity tests, Spearman rank correlation coefficients (Rho) were calculated using Systat 7.0 software. Since the response of the control groups for each toxicity test was both acceptable and consistent, the sediment toxicity test data were not normalized to control results. Rho values, corrected for ties, were determined for each toxicity test and each pollutant or pollutant class, and were compared to tables at the appropriate n value to determine the level of statistical significance associated with the observed correlation.

Weight-of-Evidence and Categorization of Sites

Toxicological, chemical, and ecological measures were combined to provide a weight-of-evidence categorization of sediment quality at each site. This approach is consistent with generally accepted methods of sediment quality assessment, such as the commonly used "sediment quality triad" described by Chapman et al. (1987). The three primary measures in the triad approach are sediment chemical analysis, toxicity testing, and benthic community analysis. All of these measures have their advantages and drawbacks, but together they can be used effectively to characterize sediment quality. In San Francisco Bay, toxicity testing was used as the primary screening tool in the first round of sampling. Stations that produced toxic samples or had been shown in previous studies to have elevated chemistry, bioaccumulation, or other measures of pollution were then resampled and analyzed for toxicity, chemistry, and, to a lesser extent, benthic community structure. Benthic community measures were originally de-emphasized because of the difficulty interpreting benthic data from San Francisco Bay, where waves of invading exotic species and extreme salinity fluctuations strongly affect benthic communities. Benthic data were eventually collected at 14 stations during the confirmation phase of the study, and these data were supplemented with laboratory bioaccumulation data from 10 stations.

Use of Threshold Values

Using the data collected in this study, stations were categorized based on chemical concentrations, the severity of biological impacts, and the completeness of sample characterization. The conceptual framework for categorizing stations is provided in the listing below. In order to categorize stations, it was necessary to define terms such as "elevated chemistry", "sample toxicity" or "degraded benthos" for a large number of samples. To be consistent, thresholds were established for this purpose. Those thresholds are defined below in the description of the first category. Toxicity thresholds were based on the reference envelope tolerance limits for amphipod tests, and t-test plus detectable difference criteria for sea urchin development tests (as defined above, because sea urchin tolerance limits were higher than could be justified based on best professional judgement of the biological significance of small

differences from controls). Benthic community degradation was defined as a Relative Benthic Index ≤ 0.30, based on the best professional judgement of the ecologists who developed the index. Elevated chemistry was defined as 6 or more chemicals exceeding ERM guidelines, a mean ERMQ above 0.5, or one or more chemicals at concentrations high enough to likely be associated with biological effects, based on best professional judgement. The mean ERMQ value of 0.5 was based on an evaluation by Long and MacDonald (in press) that indicated at least 50% of samples in a nationwide evaluation exhibited toxicity when this value was exceeded. The BPTCP has calculated mean ERMQ values using a different suite of chemicals than used by Long and MacDonald (in press); the primary differences being that Long and MacDonald (in press) used a number of individual PAHs and the DDT ERM, whereas the BPTCP used only the summary low and high molecular weight PAHs (2 values) and the DDT value of Swartz et al. (1994). When the mean ERMQ values, as calculated by the BPTCP, were compared with amphipod toxicity in the statewide BPTCP database, 62% of the samples with mean ERMQs greater than 0.5 were found to be toxic to amphipods.

These chemistry, toxicity, and benthic community threshold values were derived to allow a consistent interpretation of data from samples throughout the Region and state. It is important to note that while these threshold values were selected based on the best available information and best professional judgement of the authors, they are by nature discretionary. Chemical bioavailability varies from sample to sample, and the exact definitions of toxicity and benthic degradation depend on factors not easily analyzed in a large number of samples. Further data collection and analysis may result in the determination of different threshold values and different definitions for biological impacts. The thresholds and station characterizations used here are not intended to be absolute. They are intended to aid in the screening of data collected from a large number of locations and to support management decisions for further action. In some cases, additional studies may be undertaken to further evaluate the sites of concern identified in this Region-wide assessment. If more data become available through additional studies at selected locations, more accurate site-specific characterizations of sediment quality may result.

Weight-of-Evidence Categorization Criteria

A weight-of-evidence approach was used to combine toxicity, chemistry, and benthic community data in order to categorize stations. This categorization was intended to assist in comparisons of sites that might be considered for management activity. The following is a list of categories, followed by the criteria and threshold values used to determine how each station was grouped.

Category 1:
Stations with elevated chemistry, recurrent toxicity, and degraded benthos.

Elevated Chemistry was indicated by either:

1) a guideline (ERM) quotient mean above 0.5, indicating a mixture of pollutants, or
2) six or more chemicals having concentrations above guideline (ERM) values, or
3) one or more individual chemicals at concentrations high enough to likely be associated with biological effects, based on best professional judgement.

Recurrent toxicity was indicated when at least two samples collected at different times from a station or site were determined to be significantly toxic by any of the BPTCP toxicity test protocols.

Degraded benthos was indicated by a Relative Benthic Index score of 0.30 or less, as described above.

Category 2:
Stations with elevated chemistry, toxicity in one (of one) sampling event, and degraded benthos.

Category 3:
Stations with highly elevated sediment concentrations of chemicals cited in the San Francisco Bay Fish Advisory (mercury and PCBs).

Category 4:
Stations with elevated chemistry, and biological impact measured by either toxicity or degraded benthos (with no data available for the second biological indicator):

a. Stations with elevated chemistry, degraded benthos, and no available toxicity data.
b. Stations with elevated chemistry, recurrent toxicity and no available benthics data.
c. Stations with elevated chemistry, toxicity in a single sample and no available benthics data. (only one toxicity sample tested)

Category 5:
Stations with elevated chemistry and mixed results from biological indicators.

a. Stations with elevated chemistry, degraded benthos, and multiple toxicity tests with some toxic and some non-toxic.
b. Stations with elevated chemistry, degraded benthos, and toxicity data indicating samples were non-toxic.
c. Stations with elevated chemistry, recurrent toxicity and data indicating non-degraded benthos.
d. Stations with elevated chemistry, toxicity in a single sample and data indicating non-degraded benthos. (only one toxicity sample tested)
e. Stations with elevated chemistry, data indicating non-degraded benthos and multiple toxicity tests with some toxic and some non-toxic.

Category 6:
Stations with measured biological impact but chemistry values below thresholds.

a. Stations with recurrent toxicity, and degraded benthos, but no chemistry data available.
b. Stations with recurrent toxicity, and degraded benthos, and elevated NH_3 or H_2S* but no other elevated chemistry.
c. Stations with recurrent toxicity, and degraded benthos, but existing chemistry data has no chemicals measured at elevated concentrations.
d. Stations with a single indicator of biological effect (either recurrent toxicity or degraded benthos), but existing chemistry data has no chemicals measured at elevated concentrations.
e. Stations with a single toxic sample, but existing chemistry data has no chemicals measured at elevated concentrations

Category 7:
Stations with chemistry, toxicity, benthic degradation below thresholds.

Reference Stations
These were selected in a previous study (Hunt et al., 1998), based on location distant from sources of pollution, chemical concentrations below guideline values, fine grain size, and minimal biological impacts.

Ranking within these major categories was determined by the actual data values, for example, 20% survival would be considered worse than 55% survival, etc.

*Elevated concentrations of NH_3 or H_2S may have resulted from natural processes, or may be enhanced by human activity. They can therefore be considered either as interferences in toxicity or benthic assessments, or as anthropogenically-derived toxins. This should be considered on a site-by-site basis.

RESULTS AND DISCUSSION

The data figures and tables discussed below are presented consecutively following the Conclusions section.

Sediment Toxicity

Data Quality

The primary quality assurance criterion for toxicity testing in this study was satisfactory survival or normal larval development in laboratory controls, as described by Stephenson et al. (1993). All toxicity data presented in this report were generated in tests with satisfactory control response, indicating that any negative responses observed in test organisms were due to sample characteristics rather than initial organism health or testing conditions. Minor deviations of QA criteria were observed in tests of a number of samples, as indicated by quality assurance codes of "-3" or "-5" in Appendix E. These deviations included test solution dissolved oxygen concentrations above 100% saturation, salinity variations of more than 3 ‰, temperature variations of greater than 1° C, and precision estimates for these water quality parameters (or for pH, ammonia or sulfide) beyond the specified ranges. Quality assurance reports were delivered to the State Water Resources Control Board with every data report, and details of minor QA exceedences can be found there. None of the variations from QA criteria were expected to affect the results of any tests from which data presented here were derived. If critical management decisions must be based on data for which a "-5" code is assigned in Appendix E, the reader is advised to consult the reports on file at the SWRCB.

Toxicity Screening

Solid-Phase Amphipod Tests

Samples from throughout the region were screened for solid-phase sediment toxicity (Figure 4). Of 122 samples, 26 (21%) were significantly toxic relative to reference envelope tolerance limits (Table 13). Toxic samples were collected from a wide variety of locations, including: Stege Marsh, Carlson Creek, and Castro Cove near Richmond, Peyton Slough in Suisun Bay, Pacific Drydock, Fruitvale, Schnitzer, and San Leandro Bay near Oakland, Mayfield Slough and San Bruno Channel in the South Bay, Islais Creek and Mission Creek in San Francisco, Waldo Point in Marin County, and Marconi Cove in Tomales Bay. Most of these sites were selected for additional sampling in the confirmation phase of the study. Of note in the screening data is the extremely high ammonia value for sample 1503 from Oakland Fruitvale (Table 13). This sample had high ammonia at the beginning of the exposure, but the concentration increased by two orders of magnitude during testing due to decomposition of dead clams. Hydrogen sulfide was also elevated, and dissolved oxygen was very low at the end of the exposure. Amphipod survival was 16% in this sample.

Sea Urchin Larval Development Tests in Porewater and at the Sediment-Water Interface

Of 99 porewater samples screened for toxicity to sea urchin larvae, 31 (31%) were toxic (Figure 5). Forty-two of the 99 samples were also tested in dilutions of 50% and 25% porewater; of

these, 34 (81%) were toxic at all dilutions, seven (17%) were toxic only at full strength, and one (2%) was toxic at 100% and 50% dilution but non-toxic at 25% dilution. None were toxic in lower dilutions when the full strength porewater was non-toxic. Ammonia and/or hydrogen sulfide were above threshold toxicity limits in many samples (Table 14).

In sediment-water interface (SWI) exposures, 11 of 33 samples (33%) were toxic to sea urchin larvae (Figure 6, Table 14). Three of the toxic samples had ammonia or hydrogen sulfide above threshold values, and three of the non-toxic samples had ammonia or hydrogen sulfide above threshold values, providing evidence of the uncertainty involved in the estimation of toxicity thresholds for these (and other) compounds. Further studies are necessary to refine ammonia and sulfide toxicity thresholds, including investigation of critical exposure periods over the embryo/larval development cycle.

Toxicity at Confirmation Stations

Twelve sites, some incorporating multiple stations, were resampled in the confirmation phase, based on toxicity or chemistry data from screening analyses. Eight of the 12 confirmation stations exhibited significant toxicity to amphipods during screening, and all eight were again toxic to amphipods in the confirmation tests. Samples from five of these eight stations were also toxic to sea urchins (Figures 7 and 8). Trends along gradients from selected confirmation stations, along with TIE and other supporting information are discussed in subsequent sections.

Four sites selected for confirmation were toxic only to sea urchins in pore water during screening (Table 15a). None of these were toxic to sea urchins in confirmation phase SWI tests, but one (Central Basin) was toxic to amphipods.

The Oakland Fruitvale station that produced the sample with extremely high ammonia in the screening phase did not have elevated ammonia concentrations in the confirmation sample, and hydrogen sulfide was only slightly above the threshold value (0.1231 mg/L compared to the threshold of 0.114 mg/L). Amphipod survival in this sample was 55% (Table 15a; Appendix E). Additional samples in which ammonia or hydrogen sulfide may have influenced toxicity test results (Table 15) are not discussed in detail here, but can be investigated by comparing values for those samples (Appendix E) with threshold values (Table 1). Elevated ammonia and hydrogen sulfide may be the result of natural processes, or their concentrations may be enhanced by human activities, such as discharge of organically enriched wastewater. These compounds can therefore be considered as natural interferences in toxicity tests (if of natural origin) or as pollutants in need of management (if anthropogenically enhanced). In either case, elevated ammonia or hydrogen sulfide concentrations do not in any way preclude the presence of other chemicals in toxic concentrations.

In addition to confirmation sampling at the stations described above, a set of eleven other stations were also sampled a second time to investigate toxicity to sea urchin larvae in porewater samples rich in ammonia or hydrogen sulfide (Table 15b). Porewater samples from these eleven stations were toxic to sea urchin larvae during screening, but porewater ammonia and/or

hydrogen sulfide concentrations were above thresholds, solid-phase samples were not toxic to amphipods, and available chemistry data indicated generally low concentrations. In the second round of testing at these stations, samples were tested using SWI exposures to provide a more environmentally relevant exposure, and to minimize the effects of porewater ammonia and hydrogen sulfide (Anderson et al., 1996). Only two of the eleven were found to be toxic to sea urchin larvae exposed at the sediment-water interface (Table 15b). None of the SWI test overlying water solutions had ammonia or hydrogen sulfide above threshold values. In the two samples with SWI toxicity, the magnitude of the response was not as extreme as in the porewater samples.

While testing porewater is useful for investigating bioavailability and routes of exposure for sediment associated chemicals, sea urchin embryos do not naturally occur in porewater environments. The negatively buoyant embryos may come into contact with sediment surfaces, however, and the SWI exposures were designed to investigate effects of chemicals fluxing from polluted sediments. The SWI exposures also dilute porewater chemicals, including ammonia and hydrogen sulfide, because the exposures are initiated with clean laboratory seawater overlying the test sediments. While the SWI exposures reported here resulted in decreased toxicity, it cannot be inferred that the sediment porewaters were any less toxic in the second set of samples, or that the observed porewater toxicity was due solely to ammonia or hydrogen sulfide. In addition to considerations of environmental relevance, the SWI exposures have two other advantages. One is that intact sediment cores (rather than homogenates) can be tested, resulting in less disturbance of sediment equilibrium conditions. This allows more realistic assessments of chemical partitioning and flux to overlying water. A second advantage is that salinity adjustment is not necessary, because organisms are suspended above the sediment in overlying laboratory seawater. Salinity adjustment of porewater for toxicity testing can result in unequal dilution of porewater samples, especially in an estuary with wide salinity fluctuations, such as San Francisco Bay. In a recent study of S.F. Bay porewater toxicity, final porewater concentrations as low as 55% were necessary after salinity adjustment with hypersaline brine (Hunt et al, 1998; their Table 3), while other samples collected concurrently from other parts of the Bay could be tested at full strength.

Sediment Chemistry

Data Quality

All trace metal analyses met all quality assurance criteria as described by Stephenson et al. (1993). Many trace organic analyses had minor deviations from QA criteria, as indicated by "-5" QA codes (Appendix C). Most of these deviations involved blank responses outside of control chart guidelines, in which case the chemical concentrations measured were corrected based on blank response prior to reporting. If critical management decisions must be based on data for which a "-5" code is assigned in Appendix C, the reader is advised to consult the data QA reports on file at the SWRCB.

Chemical Mixtures

All stations analyzed had detectable levels of multiple chemicals (Appendix C). These chemical mixtures have resulted from the variety of pollutant sources and the complexity of chemical transport within San Francisco Bay (Kennish, 1998). Areas in which sediments accumulate also tend to accumulate sediment-associated chemicals. The resulting chemical mixtures affect organisms unpredictably, and no two suites of chemicals are identical, though trends in covariance can be detected. In an attempt to measure the cumulative level of chemical pollution, and to roughly gauge their probable impact on biological resources, guideline quotient values were used. These values, their derivation and use are described in the Statistical Analyses/Chemical Data part of the Methods section of this report. Also described there is the justification for using a threshold value for the mean sediment quality guideline quotient for screening the large number of samples investigated in this study. An Effects Range Median quotient value (ERMQ) ≥ 0.50 was used to screen for elevated chemical mixtures. ERMQ values were calculated only when both trace metal and organic chemicals were measured; in cases where only one or the other group of compounds were analyzed, an "na" appears in the ERMQ column of the data tables.

Of the 46 stations screened for trace metal and organic chemicals, 23 (50%) had mean ERM quotients greater than 0.50 (Table 16). ERMQ values ranged as high as 3.94 (at a Pacific Drydock station). Other stations with highly elevated levels of chemical mixtures included: Mission Creek (3.93), Peyton Slough (3.58), Stege Marsh #1 (2.70), Stege Marsh #2 (2.59), Castro Cove (2.25), San Leandro Bay Site 4 (2.01) and San Leandro Bay Site 1 (1.52; see Figure 9). Samples from each of these stations had numerous chemicals at concentrations above ERMs, some much greater than 10 times the guideline values. Only three stations had no chemicals measured above ERM values (Table 16). An additional nine stations exceeded only the ERM value for nickel, which is common in San Francisco Bay sediments and is derived primarily from geologic sources, though substantial anthropogenic sources exist.

Chemicals Found at Elevated Concentrations

Chemicals or chemical classes commonly found above guideline values included: chlordanes, PCBs, DDTs, PAHs, dieldrin, hexachlorobenzene*, chlorpyrifos*, copper, mercury, lead, selenium*, and zinc (Table 16). (*There are no ERM values currently derived for hexachlorobenzene, chlorpyrifos, or selenium, but these chemicals were often found above the 90th percentile of samples collected by the BPTCP statewide.) Detection of chlorpyrifos in sediments is worth noting because it's occurrence in sediments has not been extensively documented. A few of these chemicals or chemical classes are discussed below, because their presence was especially pervasive in Bay sediments (e.g., chlordane, PAHs) or in fish tissues (e.g., PCBs and mercury, the subjects of the health advisory for fish consumption).

Some chemicals were found at the highest sediment concentrations measured anywhere in the State by the BPTCP: arsenic and selenium (at Stege Marsh #1), copper and zinc (at Peyton Slough Upper #2), mercury (at Pt. Portrero #2), total PCBs (at Pt. Portrero #1), and total and high molecular weight PAHs (at Castro Cove). (See Tables 12, 16, and Data Appendices).

PCBs

PCBs have accumulated in the sediments of the estuary due to historic use and continuing deposition from storm water runoff. This class of chemicals is comprised of 209 compounds called congeners. Mixtures of congeners have been manufactured in the U.S. since 1929 and sold under the trade name Aroclor. These mixtures were used extensively in the U.S. prior to 1979 when their manufacture, processing, use and application was banned, except in totally enclosed applications such as transformers. PCBs were used for industrial applications requiring fluids with thermal stability, fire and oxidation resistance, and solubility in organic compounds. PCBs have proven to be extremely persistent in the environment. RMP monitoring data indicate that PCBs in the water column exceed non-promulgated U.S.EPA water quality criteria throughout the estuary. This is most probably due to resuspension from the sediments, where they provide an ongoing source to organisms in the food chain (RMP, 1997). These substances are highly lipophilic, have a high potential to accumulate in the tissues of aquatic organisms, and can represent a significant human health hazard (Moore and Walker, 1991).

Although the use of PCBs has been banned there are historic deposits in the sediment and on land. Point Portrero, at the Port of Richmond, had ten times the PCB concentration (19.9 ppm) of any other sample collected in this study (Table 17). Other stations with elevated sediment PCB concentrations include Stege Marsh, Yosemite Creek, Islais Creek, Pacific Drydock, San Leandro Bay, and Mission Creek (Figure 10). Stormwater events can mobilize PCBs deposited on land and transport them into the estuary. Recent monitoring by the RMP indicates the existence of current sources contributing to PCB loads into the South Bay from Coyote Creek. Increased monitoring is necessary to identify active sources and minimize their impacts.

Mercury

Mercury was mined in the Coast Range from the early 1800's through the mid-1900's. Initially, most of the mercury was used in the amalgamation of gold in placer and hydraulic mining operations. Mining activity introduced mercury into the San Francisco estuary system in a number of ways. Runoff from mercury mines within the region transported mercury-rich sediment into the Bay and estuary. In the Sierra, mercury was added to sediment to aid in the separation of gold from waste in placer and hydraulic mining operations. Most of this mercury ended up in the aquatic system, becoming attached to sediment particles flushing downstream. Mining of gold and silver ores exposed surrounding rock that was enriched in mercury by the same geologic processes that created the gold and silver deposits, again introducing mercury-rich sediment into the stream systems that drain into San Francisco Bay. Continuing drainage from these mines has introduced mercury and other metals into the streams that drain into the estuary (Kennish, 1998).

The estuary, therefore, has become a sink for sediments rich in mercury that are likely to be a continuing source for the bioaccumulation of mercury up the food chain. Data from the current study indicate that mercury concentrations in the estuary are elevated but highly dispersed (Table 17). There are a number of individual sites around the margins of the Bay where mercury concentrations are higher than historically elevated levels. These are usually due to past industrial practices such as the smelting of ore. Stations with highly elevated mercury

concentrations include: Point Portrero, Mission Creek, Pacific Drydock, Kirker Creek, Stege Marsh, and Castro Cove (Figure 11).

Although mining practices were historic, runoff from abandoned mines and mine tailings continue to be an ongoing source of mercury to the estuary. Data from the Sacramento River indicate that the Cache Creek drainage and the Sacramento drainage above the Feather River are major, ongoing sources to the lower watershed (personnal communication, Chris Foe, Central Valley Regional Water Quality Control Board). In the southern part of San Francisco Bay, the major ongoing source is the drainage from New Almaden mining region (SFBRWQCB, 1998). Other less significant sources include POTWs, industrial discharges and aerial deposition. Recent pollution prevention audits indicate that human waste, water supplies, laundry waste, household products, thermometers, and waste from hospitals and dental facilities are the most significant sources flowing into POTWs (SFBRWQCB, 1998). Known industrial discharges of mercury are from raw materials used in the facilities. About half the aerial deposition appears to come from global fuel combustion and the other half from local fuel combustion.

The key environmental concern about mercury in the San Francisco Bay system is the extent to which it bioaccumulates in the food chain. Bioaccumulation, in turn, is governed by the level of methyl mercury in the Bay sediment system. Methyl mercury is formed primarily by microbial activity, and only under certain physical and chemical conditions. Different forms of mercury as well as different environmental conditions may increase the rate of mercury methylation. This process must be better understood in order to regulate the current reservoir of mercury and minimize the creation of environments that may increase the rate of methylation.

Trace Metals (Other than Mercury)

A number of trace metals were measured at elevated concentrations in numerous samples (Table 16). Nickel and chromium, as previously discussed, were elevated throughout the Bay, due generally to geologic abundance and specifically to human inputs in some areas. Other metals commonly found at elevated concentrations included arsenic, copper, lead, selenium, silver, and zinc. These are derived from a variety of sources, including foundry, electroplating, etching and other industrial operations, oil refining, piping systems, POTWs, agricultural tail water, urban runoff, construction grading, mining operations, ore processing, and natural rock formations.

Metals entering aquatic systems often bind to charged particles such as clays, and thus accumulate in sediments, where they can form insoluble metal sulfides under anaerobic conditions. The sediment concentration of acid-volatile sulfide (AVS) under anaerobic conditions has been shown to affect the toxicity of some sediment-associated divalent metals, such as cadmium and nickel, and, to varying degrees, copper, lead, and zinc (DiToro et al., 1990, 1992). Only when combined concentrations of these metals exceed those of AVS (on a molar basis) will they be available for other binding phases (such as with clay or organic carbon) or for affecting organisms. Therefore, when the metals concentration (as the total molar concentration of the five metals, extracted simultaneously with AVS) is less than the AVS concentration, these metals should not be toxic in anaerobic sediments. In order to gain insights into the bioavailability and potential toxicity of trace metals in sediments from stations in San Francisco Bay, we measured SEM (simultaneously extracted metals) and AVS concentrations. SEM-AVS

data and their relationship to toxicity were interpreted cautiously, however, because surficial sediments collected in the present study were generally not anaerobic, and sediments were necessarily aerated during sampling and toxicity testing.

The SEM concentration exceeded that of AVS in sediments from stations at Peyton Slough, San Leandro Bay and Paradise Cove (Table 18). Samples from the two Peyton Slough stations with SEM > AVS were toxic to amphipods, with one of the two being toxic to sea urchin larvae in SWI tests. The San Leandro Bay Site 4 sample had SEM > AVS and was toxic to amphipods and sea urchins. In samples from three other stations in which SEM exceeded AVS by lesser amounts (Paradise Cove, San Leandro Bay 5, and San Leandro Bay 6), significant toxicity was not observed. Metals at these stations may have been bound by organic carbon and/or may have been at concentrations below toxic levels. For example, the SEM concentration at Paradise Cove was 2 to 20 times lower than SEM concentrations at San Leandro Bay and Peyton Slough stations (Table 18).

Since metals not bound to AVS may bind to other sediment constituents, such as organic carbon and clay particles, SEM-AVS data are generally used as evidence to eliminate metals as causes of toxicity when SEM is less than AVS. Among samples in which SEM was less than AVS (Mission Creek, Islais Creek, Pacific Drydock, and other San Leandro Bay stations), there was wide variation in toxicity, from 100% amphipod mortality to high levels of survival and normal development (Table 18). Concentrations of organic chemicals were elevated at all of these stations, as often were concentrations of hydrogen sulfide and ammonia, and these compounds could have contributed to toxicity in the absence of trace metal effects. The relationships between trace metal concentrations and biological effects are discussed further in subsequent sections dealing with correlations, TIEs and gradient studies.

Chlordanes

As in previous BPTCP statewide monitoring studies (Fairey et al., 1996; Anderson et al., 1997; Anderson et al., 1998), and in San Francisco Bay Regional Monitoring Program studies (RMP, 1997) total chlordane was one of the pesticides most commonly measured at elevated concentrations (Figure 12). Total chlordane is the summation of the major constituents of technical grade chlordane and its metabolites (in this case cis- and trans-chlordane, oxychlordane, and cis- and trans-nonachlor; Appendix C). These comprise a group of nonsystemic stomach and contact insecticides which until the mid 1970's had been used extensively in home and agricultural applications. Although the use of this compound was discontinued in this country due to it's widespread occurrence, biomagnification through the foodchain, and persistence in non-target systems, chlordane continues to occur in aquatic ecosystems. Due to their limited water solubility, chlordane compounds tend to bind to organic carbon and settle out of the water column, accumulating in sediments (Wilcok et al., 1993). Stations with the highest measured concentrations of total chlordanes included Pacific Drydock, Mission Creek, and San Leandro Bay (Table 16). These sites receive urban runoff from creeks and storm drains.

DDTs

DDT and its metabolites are a class of relatively water-insoluble organo-chlorine compounds which tend to bind to organic particulates and thus accumulate in sediments. Concentrations of these compounds have generally declined in aquatic ecosystems since they were banned for most insecticide applications in 1972, although concentrations of some DDT metabolites have increased. Like chlordane and dieldrin, it is persistent in sediments and may be of significant environmental concern at higher concentrations (Hoke et al., 1994; Swartz et al., 1994). Sites with elevated concentrations of DDT compounds (primarily p'p' DDE) included Stege Marsh, Islais Creek, Pacific Drydock, Peyton Slough, and San Leandro Bay (Table 16). One important source of DDT in the Bay, Lauritzen Channel in Richmond Harbor, was not sampled in this study. This site had extremely elevated levels of DDT, and was recently cleaned up under the federal CERCLA program, with the U.S.EPA acting as the lead agency.

PAHs

Polycyclic Aromatic Hydrocarbons (PAHs) are base-neutral organic compounds that are components of crude and refined petroleum products. They are also produced by incomplete combustion of hydrocarbons. These compounds are common components of contaminated sediments and are toxic to infaunal invertebrates (Eisler, 1987; Neff, 1979; Neff and Anderson, 1981), in particular amphipods (Swartz et al., 1995). Due to their similar modes of toxicity, individual PAHs are combined into low and high molecular weight groups.

A number of stations had measured concentrations of PAH compounds above guideline values, with the highest concentrations at Castro Cove and Pacific Drydock. Castro Cove had highly elevated concentrations of individual PAH compounds (such as dibenzo(a)anthracene) and total high molecular weight PAHs. Surficial sediments from this site contain a distinct petrogenic profile dominated by alkylated high molecular weight PAHs. This is in contrast to the pyrogenic (derived from combustion) profile which is the prevalent PAH contamination profile in San Francisco Bay. The high relative abundance of alkylated naphthalenes, the dominance of anthracene over phenanthrene, and the abundance of alkylated high molecular weight PAHs show Castro Cove to be unique. The low water solubility of the dominant PAH compounds suggests a local source and/or historical release (Newman, 1998).

Bioaccumulation Tests

Data Quality

The bioaccumulation data presented here should be interpreted with caution. Only one laboratory replicate of each sample was analyzed, so the precision of quantitative estimates of chemical bioaccumulation is unknown. Clams were exposed to additional laboratory replicates of each sample, but the storm event of February 2, 1998, cut power and access to the laboratory at which the samples were held in frozen storage, and the additional replicates thawed before power could be restored. In addition, control survival of clams over the 28-day exposure in one set of tests was less than recommended in the test protocol (ASTM, 1996). Because of these limitations, the data were not analyzed statistically, but rather were evaluated quantitatively only to determine whether they were at least ten times greater than comparable control values, or whether they exceeded EPA or NAS guidelines. While bioaccumulation data were not listed in the final weight-of-

evidence table. they provided a qualitatively indication of which sediment associated contaminants were bioavailable. Information about bioavailability should be considered along with sediment chemical concentrations to evaluate risks of biological impacts at test sites.

Chemicals Detected in Tissues of Exposed Clams

Chemicals or chemical classes found in exposed clam tissues at concentrations at least 10 times higher than in controls included copper, lead, total chlordanes, total DDTs, dieldrin, total PCBs, and low and high molecular weight PAHs (Table 19). Total DDT was found to exceed the NAS guideline for protection of wildlife in one sample (from Stege Marsh). US EPA screening levels for protection of human health were exceeded for total PCBs at Mission Creek, Pacific Drydock, San Leandro Bay and Stege Marsh. PCBs and mercury were two chemicals cited in the California Office of Environmental Health Hazard Assessment health advisory for fish consumption in San Francisco Bay and the Delta. While PCBs were highly accumulated in clam tissues after laboratory exposures, mercury never accumulated by more than a factor of about two over controls in these tests (Table 19a). There are a number of possible reasons for this, including: 1) clams were collected from Tomales Bay, which receives drainage from mercury mine operations, and control tissue levels may have been elevated prior to testing; 2) bioaccumulation rates are governed by organism physiology, and bivalves may not accumulate mercury as efficiently as other resident biota; and 3) methyl mercury is the most bioavailable form, and the degree of methylation in test sediments was unknown.

Stations with Samples Exhibiting Bioaccumulation

Bioaccumulation was measured in samples from 10 stations, including a reference station at Paradise Cove. Clams exposed to samples from this reference station and two other stations, Islais Creek and Warm Water Cove, did not accumulate any chemicals by factors ≥ 10 times control values, nor did they exceed NAS or EPA guidelines in these tests (Table 19d). Stations with chemicals that accumulated substantially in test clams included: Peyton Slough (copper and lead), Mission Creek (PCBs), Stege Marsh # 1 (PCBs), San Leandro Bay (lead and chlordane), Pacific Drydock (lead, PCBs, low MW, high MW and total PAHs), Stege Marsh # 2 (DDTs, PCBs, low MW, high MW and total PAHs), and Stege Marsh # 3 (DDTs, Dieldrin, PCBs, high MW and total PAHs). These data can be interpreted qualitatively to indicate that chemicals in these samples were bioavailable and were capable of accumulating in tissues of exposed organisms.

Benthic Community Structure

Data Quality

All benthic community analyses followed the quality control procedures described in the Methods section of this report and in the BPTCP QAPP (Stephenson et al., 1994). All resulting data met the quality assurance criteria described in Stephenson et al. (1994).

Benthic Community Structure at Sampled Stations

Benthic communities were classified as degraded (Relative benthic Index[RBI] ≤ 0.30) in samples from five of the 22 stations studied (23%; Table 20). No living organisms were found in samples from either of the Stege Marsh stations. Two stations in Islais Creek had degraded

benthos, determined primarily by the absence of amphipods or mollusks. One station in Mission Creek had an RBI of 0.00, based primarily on the dominance of pollution tolerant Oligochaetes (Appendix F). This and a sample from another Mission Creek station had no mollusks and only one amphipod between them.

Stations with RBI values above 0.30 but still relatively low (transitional category: 0.31 to 0.60) included Peyton Slough (no mollusks), and the San Pablo Bay Island # 1 reference station (no amphipods). Other characterizations of benthic communities at this San Pablo Bay station have indicated possible benthic impacts (RMP, 1997). Most of the samples from San Leandro Bay were characterized as non-impacted, though two of these were right at the RBI cutoff value of 0.60 (Table 20). (Limitations regarding the use of RBI values and associated thresholds are considered in the Methods section.) The other two sediment reference sites, Paradise Cove and North South Bay, had transitional benthic communities, with no mollusks present at Paradise Cove. These reference sites have been shown to have relatively low pollutant concentrations (Hunt et al., 1998), and the characterizations of benthic communities at these sites as transitional (moderately impacted) reflects the difficulty in assessing pollution impacts on benthic communities in an area subject to high levels of physical and biological disturbance. These difficulties are important considerations in interpreting the results of benthic analyses here. Many of the stations sampled are subjected to either very low or very high salinity, grain size and TOC fluctuations, and severe biological impacts from exotic species invasions.

Relationships between Sediment Chemistry and Biological Effects

Principal Components Analysis and Correlations

Many chemical classes have affinities for sediment particles, and chemical concentrations in sediments tend to covary. Principal Components Analysis (PCA) of 46 samples for which synoptic chemical and biological data were collected found a number of incidences of chemical covariance. Many of these covarying groups of chemicals (PCA factors) were negatively associated with biological indicators (i.e., as concentrations increased, normal biological function decreased).

Amphipod survival in toxicity tests was negatively associated with concentrations of a number of covarying chemicals (Table 21). Of these, chlordanes and the PAH 2-methylnaphthalene were also found to be significantly negatively correlated with amphipod toxicity in Spearman Rank (univariate) correlations, and had at least some samples with concentrations above ERM guideline values. Principal Components Analysis also indicated that amphipod survival was negatively associated with the number of chemicals exceeding ERM guidelines and the mean ERM quotient.

No significant associations were found between chemical concentrations and sea urchin development in undiluted porewater, probably because sea urchin development was not affected by some samples and was almost completely inhibited by others, and the lack of intermediate response provided insufficient resolution to elicit significant correlations. Sea urchin development in 50% and 25% porewater dilutions were negatively associated with a number of

metals in PCA, and four of these metals (Cd, Cu, Ag, and Zn) were found at concentrations above ERMs, indicating potential for causal associations with biological effects. Summary ERM quotients for trace metals and mean ERM quotient values also were negatively associated with sea urchin embryo/larval development in the PCA. While these factors were associated with toxicity in PCA, and concentrations were above those associated with biological effects in other studies (e.g., Long et al., 1995), none were significantly correlated with toxicity in univariate correlations, as will be discussed at the end of this section.

Sea urchin embryo/larval development at the sediment-water interface (SWI) was negatively associated with concentrations of a number of chemicals in PCA. Of these, a number of metals were both correlated with toxicity and had concentrations above guideline values (Table 21). These included cadmium, copper and zinc. The number of chemicals exceeding ERM guideline values was also correlated with SWI toxicity.

The Relative Benthic Index was associated with chemical mixtures in PCA. The summary metal ERM quotient and the mean ERM quotient were both negatively associated with benthic community structure indices (Table 21).

Toxicity Identification Evaluations

The Toxicity Identification Evaluations (TIEs) reported here were conducted by exposing sea urchin embryos to samples of sediment porewater and to test solutions produced by chemical and physical fractionation of porewater. Sea urchin embryos are sensitive indicators of a number of toxins, and are known to be sensitive to trace metals, notably copper, silver, and zinc (Bay et al., 1993). TIE concepts and techniques are described in the Methods section.

Guadalupe Slough

In the initial test of porewater from Guadalupe Slough (Station 21041, Figure 2d), no embryos developed normally in 100% porewater, and 98% developed normally in 25% porewater. TIE treatments used 100%, 50% and 25% porewater, plus blanks for each treatment. The control and 25% porewater treatment in the TIE baseline test had high rates of normal development, while toxicity increased from 50% to 100% porewater. (The baseline test is conducted with the original unmanipulated sample for comparison with concurrently tested TIE treatments.) Of all the TIE treatments, only addition of EDTA substantially reduced sample toxicity (Table 22). Since EDTA binds divalent cations, this indicates that trace metals were responsible for the observed toxicity. Filtration through the C-8 column slightly reduced toxicity, more so than did simple filtration on glass fiber filters, but compounds eluted from the column were not toxic, indicating that organic compounds were marginally, if at all, responsible for sample toxicity. Of the trace metals measured in Guadalupe Slough samples, only nickel was above the ERM guideline value (Table 16). However, since chemical analyses were conducted only on bulk-phase sediment, and no SEM-AVS analyses were conducted at this site, no information is available to predict metal bioavailability. Analysis of porewater metal concentrations would allow comparisons with known effects concentrations to develop a toxic units approach to identify chemicals causing toxicity.

Peyton Slough

The Peyton Slough porewater sample was highly toxic, with toxicity being detected in porewater diluted to 13% strength in the initial test (Table 23). The TIE tested concentrations up to 15% porewater, which elicited 0% normal sea urchin embryo/larval development in the baseline test. Toxicity was reduced by addition of EDTA, addition of STS, and filtration with both glass fiber filters and C-8 columns. Since the C-8 column eluate was not toxic, physical filtration rather than removal of organic compounds was apparently responsible for toxicity reduction in this treatment. Both EDTA and STS have been shown to strongly remove toxicity due to copper, cadmium and mercury (Hockett and Mount, 1996). The filtration process used in this TIE trapped clay and other sediment particles on filters through which remaining sample passed, and charged clay particles are effective at removing ionic metals from solution. Chemical analyses of bulk-phase sediment from Peyton Slough indicated high concentrations of copper, zinc, cadmium and other metals (Table 16), though SEM-AVS measurements indicated that these metals might not be bioavailable in this particular sample (Table 18). As mentioned previously, metal binding by AVS occurs primarily in anaerobic sediments, and the SEM-AVS relationship may not apply to the surficial sediments evaluated here. The combined evidence indicates that copper, zinc, or combined metal concentrations were responsible for sediment toxicity at Peyton Slough. Analysis of porewater metals would help clarify causes of toxicity at this site.

Stege Marsh

Since Stege Marsh samples were highly toxic to sea urchins in SWI exposures (Table 14), overlying water in SWI exposures was treated with EDTA to investigate whether toxicity was due to fluxes of trace metals from these solid-phase samples. Samples from Stege Marsh stations # 1 and # 2 were highly toxic, and EDTA addition did not reduce toxicity. Sea urchin normal development was 19% in the SWI exposure of the Stege Marsh # 3 sample. This toxicity was partially reduced by EDTA addition (Table 24). EDTA additions in themselves were not toxic, as indicated by control responses. This evaluation was not a comprehensive Phase I TIE, and few alternative hypotheses regarding causes of toxicity could be ruled out by the resulting data. The high toxicity of the Stege Marsh # 3 sample may have been at least partially due to divalent cationic trace metals (such as copper, zinc, etc.). The even higher toxicity of samples from Stege Marsh # 1 and # 2 were due either to compounds other than divalent cationic trace metals, to other compounds and divalent cationic trace metals, or at least partially to divalent cationic metals that existed at concentrations high enough to overwhelm the binding capacity of the EDTA. The large number of chemicals above guideline values at these sites (Table 16) indicates the presence of elevated chemical mixtures that may overwhelm initial TIE attempts to resolve causes of toxicity. Identification of causal agents could be further explored by conducting the full suite of TIE manipulations on dilutions of overlying water (or porewater), with sequential treatments to isolate multiple chemicals. The extremes of salinity and pH at this site would also need to be accounted for in any assessments of porewater toxicity.

Gradient Studies

Sampling along geographical transects away from presumed highly polluted sites allows investigations of co-occurring trends in sediment chemistry and biological effects. It also provides opportunities for rough estimation of the spatial extent of pollution. Three sites were sampled in limited gradient studies during the confirmation phase of this project: Mission Creek, Islais Creek, and Peyton Slough. All three were narrow channels, and each gradient consisted of three stations originating at the most inland station (thought to be most highly polluted) and ending near the mouth of the channel as it approached the Bay. By definition, these study designs assumed declining chemical concentrations away from the most polluted site. However, in the complex depositional environment of San Francisco Bay, multiple chemicals from multiple sources often confound delineation of simple gradients, and trends in biological effects can be difficult to resolve with respect to a given suite of measured chemicals. The likely presence of unmeasured chemicals further complicates interpretation of gradient study results. An additional limitation is that samples were collected from only three stations along each gradient, the bare minimum for resolution of linear relationships. Trends in chemical concentrations and biological effects are discussed below.

Castro Cove

An extensive gradient study was conducted at Castro Cove in an earlier phase of the BPTCP. The results of those studies have been presented by Spies et al. (1993) and Flegal et al. (1994), and are not discussed further here.

Mission Creek

During the screening phase of this project, two stations were sampled in Mission Creek, one near the upper end and one near the mouth. The upper station is near a combined sewage overflow (CSO) and is closer to sources of urban runoff. The upper station was more toxic than the downstream station to both sea urchins and amphipods. The upper station also had higher levels of multiple pollutants, as indicated by the mean ERM quotients (Table 25).

Two years passed between screening and confirmation sampling, and chemical concentrations at the upper Mission Creek station increased substantially over this period (Table 25). Mean ERM quotients increased eightfold, and metals, PCBs, and total chlordanes were 5 to 15 times their initial concentrations. Since PCBs and chlordanes have been banned for more than twenty years, these increases were likely the result of resuspension of polluted sediments during storm events, or transport of contaminated particles in urban runoff and sewage overflow.

During the confirmation phase, three stations were sampled: one at the same upper site, one midway down the channel, and one near the mouth (more than 200 m from the original creek mouth station (Figures 7f and 8f). Concentrations of a number of chemicals, including PCBs, chlordanes, lead, mercury, silver and zinc, decreased downstream, as did the mean ERM quotient. Sea urchin SWI tests showed greatest toxicity at the upper station, with insignificant toxicity at the two stations further downstream. The upper station was most highly toxic to amphipods, with toxicity decreasing downstream. Downstream sites were

less toxic even though these lower stations had 100% fine grained sediment. Hydrogen sulfide may have influenced this trend in toxicity; the upper station was above threshold values for both sea urchins and amphipods, while neither lower station was (Tables 1 and 25). The upstream site sample also had low dissolved oxygen concentrations during amphipod testing. The relative benthic index (RBI) also increased downstream, with RBI values increasing from degraded (0.00), through transitional (0.34), to undegraded (0.65; Table 25). This limited data would indicate that measures of biological effects responded predictably to pollutants, including hydrogen sulfide. The high hydrogen sulfide levels indicate a potential for dissolved oxygen depletion, which may also have played a role in benthic impacts (though we have no direct evidence of this).

Islais Creek

Islais Creek has also been influenced by a CSO outfall near the upper sampling station, and receives urban runoff. In Islais Creek, the highest concentrations of multiple chemicals (ERMQ = 1.2), the highest concentrations of some classes of compounds (chlordanes, PCBs, low MW PAHs), the highest concentrations of ammonia and hydrogen sulfide, and the highest toxicity to both sea urchins and amphipods occurred at the upper station (Table 26). Chemical concentrations decreased by a factor of about two from the upper station to the two lower stations, but chlordanes and PCBs remained above guideline values, and ammonia remained above the sea urchin toxicity threshold. Amphipod survival, the relative benthic index, and sea urchin normal development were higher at the two downstream stations. The data describe a clear distinction between the upper site and the two lower sites, rather than a gradual linear transition along the gradient.

Peyton Slough

As discussed above, TIE results suggested that sea urchin larval development was inhibited by sediment trace metals at Peyton Slough. There was a clear copper gradient at this site, with the upper station having highly elevated copper concentrations (Table 27). Other metals, especially zinc, were also much higher at the upstream site, while the two lower stations were similar to each other with lower metals concentrations. While hydrogen sulfide was below the thresholds for sea urchins in the SWI tests, ammonia was above the sea urchin threshold at the mid-gradient site. Upper and mid-gradient samples were highly toxic to sea urchins, due presumably to copper and ammonia, respectively, while the furthest downstream station was non-toxic. Salinity was a complicating factor at Peyton Slough, since the ambient salinity of approximately 2‰ was well below the test salinity of 34‰. Salinity adjustment during testing presumably had a significant effect on metal bioavailability.

The amphipod test results indicate a more complicated situation. In the first survey, the upper Peyton Slough site was highly toxic to amphipods, with only 1% survival (Table 27b). The upper site was also significantly toxic in the second survey (69% survival), when it was sampled concurrently with mid and end gradient sites. However, the two downstream sites were more toxic than the upper site, even though they had generally lower concentrations of anthropogenic chemicals. The greatest toxicity was observed in the furthest downstream site, which had hydrogen sulfide concentrations slightly above threshold levels. Both

downstream stations had 100% fine grained sediment, while the upstream station was mostly sand. Hydrometer analysis of grain size at the lower two stations indicated mostly fine silts and clays, with 35% and 45% clay at the mid and lower gradient stations, respectively. High clay content may negatively affect survival of *Eohaustorius*, though this has not been clearly demonstrated. The Relative Benthic Index characterized all three stations as transitional. However, the mid-gradient station, with very high TOC, also had very high numbers of amphipods of the genus *Corophium* (Appendix F), indicating that factors other than sediment contamination may have affected benthic communities.

The sea urchin test was probably responding to trace metals along the gradient, with high ammonia causing additional toxicity at the mid-gradient station, while amphipods and resident fauna may have been more strongly influenced by other factors such as high clay content, TOC, and/or hydrogen sulfide.

Weight-Of-Evidence Categorization Of Stations

Data Available for Station Characterization

The stations investigated in this study have been categorized using available chemical, toxicological and benthic community data to indicate the relative degree of pollution observed. This characterization relies primarily on data from the BPTCP screening and confirmation surveys, with some additional data from earlier BPTCP-funded studies, especially the Pilot Regional Monitoring Program (PRMP). A substantial set of additional data on sediment and water quality in San Francisco Bay has been collected by the Regional Monitoring Program (RMP). RMP data was used only peripherally in the following station characterization, however, because of the different objectives and sampling strategy of that program. Rather than attempting to specifically identify toxic hot spots, as is an objective of the BPTCP, the RMP has focused on repeated monitoring of a consistent set of stations selected to characterize general Bay-wide conditions. Interested readers are encouraged to compare data presented here with that generated by the RMP (e.g., RMP, 1997) to get a broader perspective on the overall status of environmental quality in San Francisco Bay.

The following characterization was based on a weight-of-evidence approach described in the Methods section of this report. That section also discusses the use and limitation of threshold values used to characterize significant toxicity, elevated chemistry, and degraded benthos.

Station Characterization

Stations at three sites could be categorized as significantly polluted based on all three triad indicators: toxicity, elevated chemistry, and degraded benthic communities. These sites were Stege Marsh, Mission Creek and Islais Creek (Table 28). Two stations at Stege Marsh (# 1 and # 2) were highly toxic to both test species, had highly elevated concentrations of numerous chemicals and had no living benthic organisms. The third Stege Marsh station had similarly elevated chemistry and high toxicity, but was not sampled to characterize the benthos. The site is unusual in that some sediment porewater had extremely low pH (pH < 4), and much of the area is high intertidal marsh subject to elevated salinity.

Mission Creek and Islais Creek had similar pollution profiles: elevated chemistry, often, but not always associated with elevated ammonia and/or hydrogen sulfide; high toxicity to both species, and degraded benthos (Table 28). These sites are described in more detail in the Gradient Studies section above.

Three stations were placed in the third category for highly elevated concentrations of mercury and/or PCBs, two chemicals identified in the fish consumption advisory (Table 28). Of these, Point Portrero is notable for having the highest concentrations of both mercury and PCBs sampled in the Bay during this study (Table 17, see data for both Pt. Portrero stations, 1 and 2). The PCB concentrations at this site were 110 times the ERM value.

A number of stations had significant toxicity and elevated chemistry (Category IV, Table 28). These included Pacific Drydock, Castro Cove, Peyton Slough, San Leandro Bay Site 1, Central Basin along the San Francisco waterfront, and the Fruitvale station in Oakland Harbor. Many of the stations have been discussed in more detail in the preceding sections. Pacific Drydock sediments were likely affected by industrial and storm water inputs; Castro Cove had highly elevated concentrations of various PAHs with a unique chemical signature; Peyton Slough had highly elevated trace metals, especially copper, that were potentially responsible for toxicity in a TIE, and were investigated in a gradient study; San Leandro Bay has been sampled at 7 stations, a number of which showed some pollution impacts warranting further investigation; toxicity at Central Basin may have been related to ammonia or sulfide, though 8 chemicals there exceeded ERM values; and Oakland Fruitvale, which is also influenced by a storm drain, and had toxicity coincident with extremely high ammonia in one survey, but also had toxicity without elevated ammonia in a second survey.

Many of the San Leandro Bay stations were placed in the fifth category; they had elevated chemistry and significant toxicity, but benthic communities appeared to be relatively un-degraded (Tables 20 and 28). Samples contained numerous amphipods, mollusks, and polychaetes. Many of the amphipods identified from these samples were of the genus *Grandidierella*, which apparently has some ability to adapt to pollution stress (Swartz et al., 1994).

The remaining stations are listed in rough order of decreasing pollution, according to the categorization criteria (Table 28; see also the Methods section). A number of stations had all available chemistry, toxicity and benthic community measures below thresholds, indicating low probability of pollution impacts. Reference site stations generally had low chemistry, low toxicity, and transitional benthos (Category VIII, Table 28). As mentioned earlier, the San Pablo Bay Island # 1 reference site did have a toxic sample and a low RBI value in one survey, exemplifying the fact that sites used in determining reference envelope toxicity tolerance limits were not pristine (Hunt et al., 1998).

CONCLUSIONS

Study Limitations

One of the primary objectives of this study was to screen a large number of sites in San Francisco Bay for indications of degraded sediment quality, and then to confirm initial findings by subsequent use of the sediment quality triad suite of toxicity, chemistry, and benthic community analyses. While this was an effective way to focus attention on the most highly polluted sites sampled, the large scope of the surveys limited opportunities to intensively investigate each site. While data from some sites clearly indicated severe pollution, additional studies are required to understand relationships between chemistry and biological effects at many stations. The large number of stations were categorized based on a number of generalized assumptions inherent in the threshold values used to differentiate sites. Data reduction for purposes of site comparison limited the level of detail to which interpretations were made. Readers interested in specific sites are encouraged to examine the data presented in the Appendices for additional detailed information that might be brought to bear on site assessments.

Funding constraints precluded conducting all analyses at all sites. Many samples that were not acutely toxic during screening were not further investigated. Many chemicals are known to be chronically toxic at concentrations that might not elicit acute toxicity in the screening tests used in this program. Also, food web biomagnification of many chemicals leads to effects on higher organisms that cannot be predicted with toxicity tests. The fact that a site was not acutely toxic should not imply that pollution problems did not exist. Logistical constraints also limited the list of chemicals for which analyses were conducted. A great number of anthropogenic chemicals are known to be present in the environment, of which the list of approximately 140 chemicals measured in the present study is just a fraction. Literature comparisons and sediment quality guideline values were generally used to screen for chemical potential to induce biological effects, even though bioavailability is sample-specific. This limitation is discussed in detail in the Methods section, and is mentioned below. The use and interpretation of benthic community analyses was impeded by the uncertainties regarding characterizations of "normal" and "degraded" assemblages in San Francisco Bay. As mentioned before, extreme seasonal salinity changes and ecological instability from successive waves of invading species make Bay benthos difficult to characterize. Finally, bioaccumulation data should be considered in a qualitative rather than quantitative sense, due to the lack of replication, as described above.

No attempts were made to determine the spatial extent of pollution, either vertically or horizontally. This information will be necessary in many cases to support management decisions.

Completion of Study Objectives

After screening 127 stations from throughout the Bay area, and returning to 12 of those for more intensive analysis during the confirmation stage, this study was successful in identifying several locations that could be described as highly polluted. The study also indicated that 21% of the samples tested were toxic to amphipods, 31% had porewater toxic

to sea urchin embryos, and 33% were toxic to sea urchin embryos exposed at the sediment-water interface. Preliminary statistical analyses indicated a number of chemicals that were correlated with biological effects, suggesting hypotheses for further investigations of the causes of sediment toxicity in San Francisco Bay.

Sites Demonstrating the Highest Levels of Pollution and Biological Effects

A number of sites had numerous chemicals with concentrations above guideline values, high concentrations of chemical mixtures, and significant biological effects. Benthic assessments were not conducted at all of these sites (and were difficult to interpret at many), so categorization and prioritization depended on the magnitude of concentrations and effects. The sites exhibiting highest chemical concentrations and most severe biological effects included: Stege Marsh, Mission Creek, Islais Creek, Point Portrero (notable for extremely high PCB and mercury concentrations), Pacific Drydock, Castro Cove, Peyton Slough, and San Leandro Bay. Many of these sites were influenced by multiple pollutant sources, including local industrial activities, urban runoff through creeks and storm drains, sewage overflows, and distant sources from which chemicals were transported through complex hydrologic and atmospheric processes.

Chemicals Of Concern

Chemicals Identified in the Fish Advisory

Mercury and total PCBs were identified in the California Office of Environmental Health Hazard Assessment health advisory on consuming fish caught in San Francisco Bay and the Delta. These chemicals were found at elevated concentrations in a number of sediment samples analyzed in this study (Table 17). PCBs, but not mercury, were accumulated to high levels in clams exposed to 6 of 10 sediment samples tested (Table 19). Mercury, but not PCBs, was found to be correlated with toxicity to sea urchins in sediment-water interface exposures (Table 21). The relationship between sediment concentrations of these chemicals and their presence in fish tissues depends on complex geochemical and ecological mechanisms that were not evaluated in this study. However, the historical nature of PCB and mercury use, and their presence at elevated concentrations in sediment, indicate that Bay sediments are a likely current source of these chemicals of concern for human health.

Chemicals Correlated with Biological Effects

All biological indicators in this study showed negative covariance (increasing concentration and decreasing biological function) with sediment quality guideline quotient means, number of chemicals exceeding guideline values, or both, in Principal Components Analyses (PCA). Chemicals identified by PCA that also exceeded guideline values and were significantly correlated with decreases in biological indicators included: total chlordanes and 2-methylnaphthalene (with amphipod toxicity); cadmium, copper, silver, and zinc (with sea urchin porewater toxicity); and cadmium, copper, and zinc (with sea urchin SWI toxicity; Table 21).

Chemicals Elevated Above Sediment Quality Guidelines

Sediment quality guidelines, as described in the Methods section, are chemical concentrations derived empirically from a large number of studies nationwide, and allow simple comparisons to concentrations shown to be associated with biological effects. This comparison is useful for perspective, but does not necessarily indicate that chemicals with concentrations above guideline values are responsible for any observed impacts. Only site-specific intensive investigations, using TIEs and other methods, can be used to determine causal relationships.

In the present study, numerous chemicals were found at concentrations exceeding guideline (ERM) values (Table 16). Of these, chlordanes, PCBs, DDTs, PAHs, dieldrin, copper, mercury, lead and zinc were commonly found above ERM values. In addition, hexachlorobenzene and chlorpyrifos, for which no ERM values are available, were often found at concentrations above the 90th percentile of the statewide BPTCP database. Overall concentrations of chemical mixtures were high at many sites, with 9 sites having mean ERM quotients above the 95th percentile of the statewide distribution.

Chemicals Elevated in Animal Tissues

Nine chemicals or chemical classes were found to bioaccumulate to elevated levels in clams exposed to 10 samples from the Bay: copper, lead, total chlordanes, total DDTs, dieldrin, total PCBs, low MW PAHs, high MW PAHs, and total PAHs. The identification of these chemicals was dependent on the particular samples tested, the physiology of the clam *Macoma nasuta*, and the 28-day exposure period of the laboratory tests. These results can be used qualitatively to indicate that many sediment-associated chemicals evaluated in this study were bioavailable, and that Bay sediments may be a source for many chemicals to which biological resources are exposed.

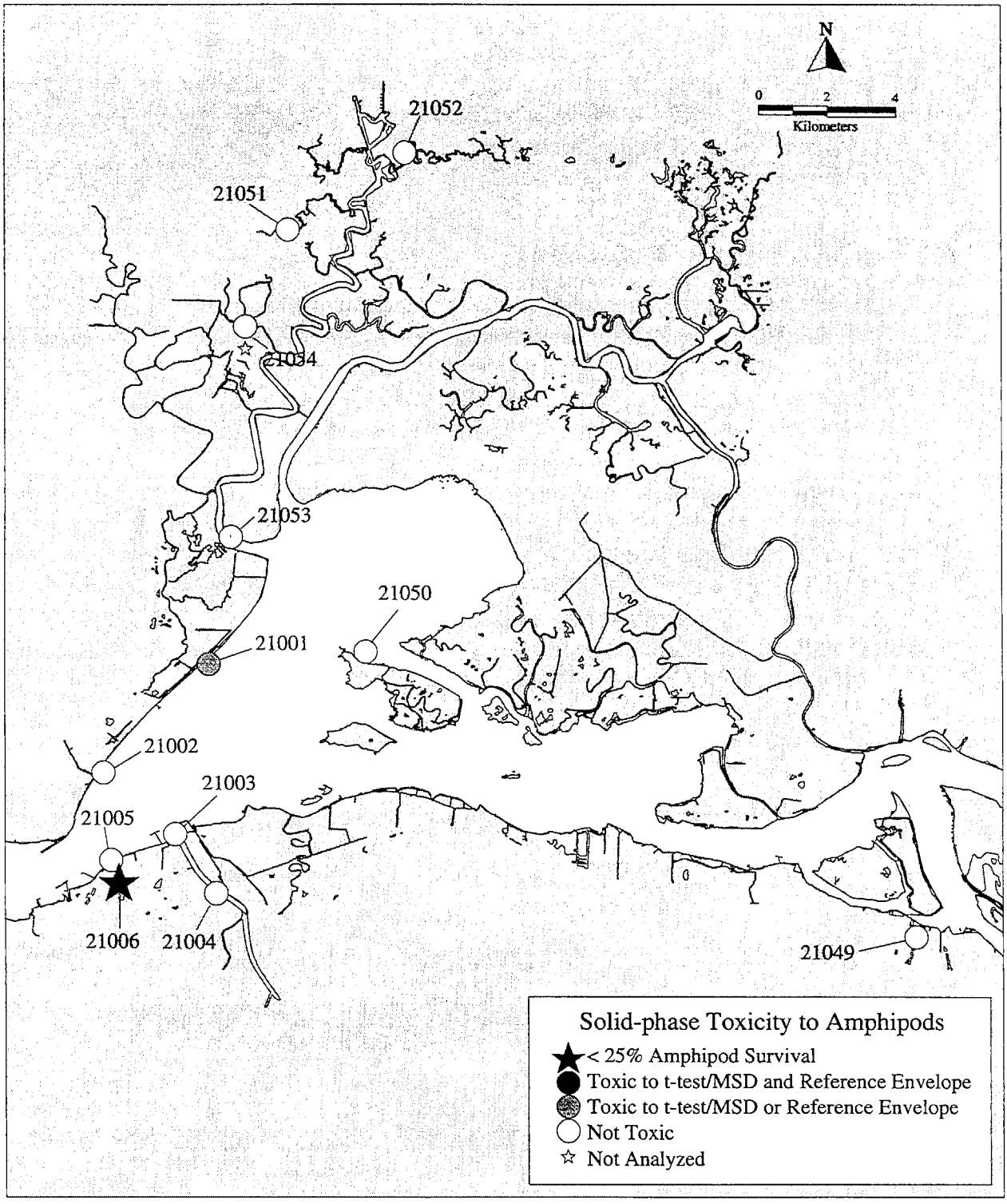

Figure 4a. Results of Amphipod Toxicity Screening for Stations in Suisun Bay.

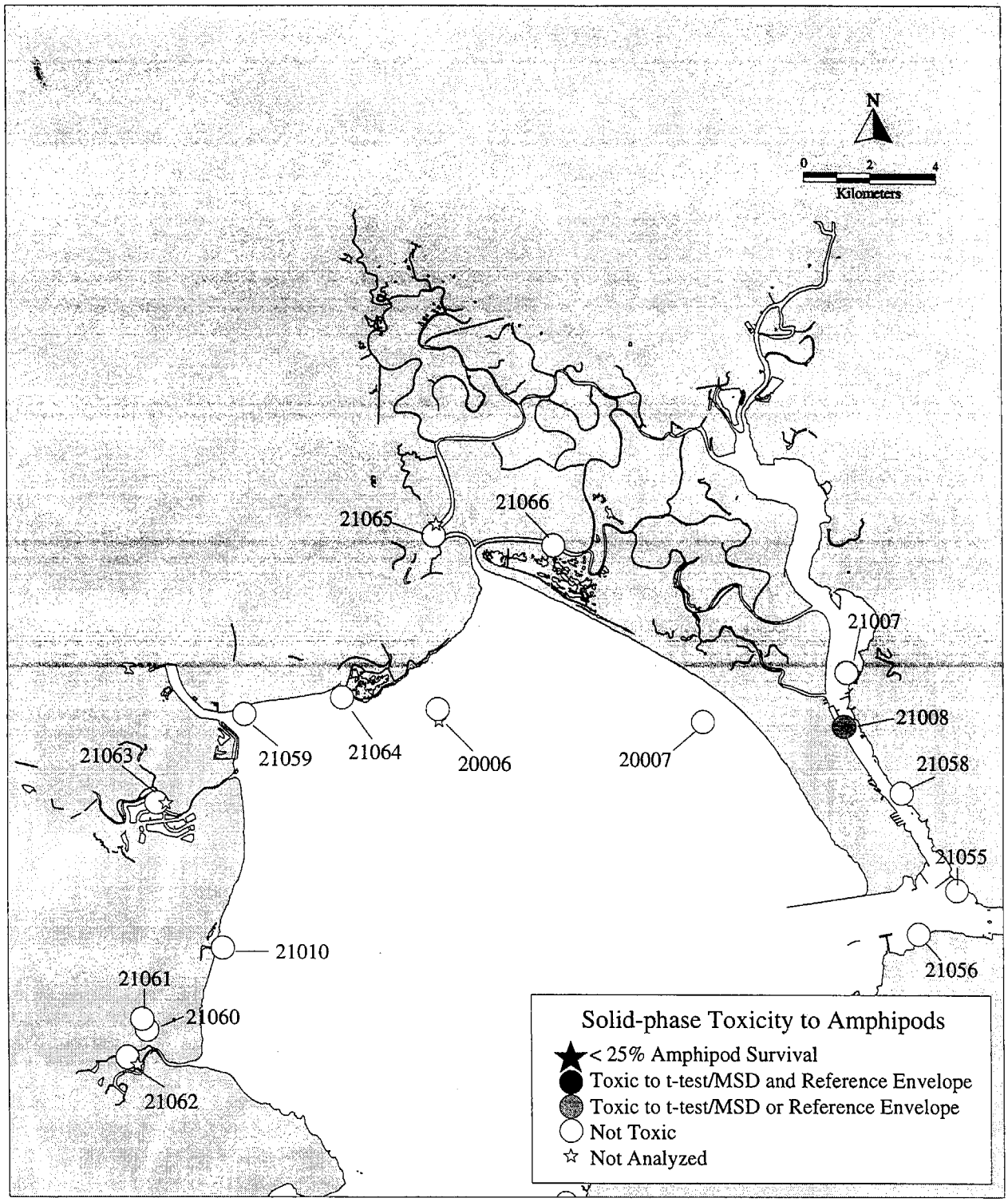

Figure 4b. Results of Amphipod Toxicity Screening for Stations in San Pablo Bay.

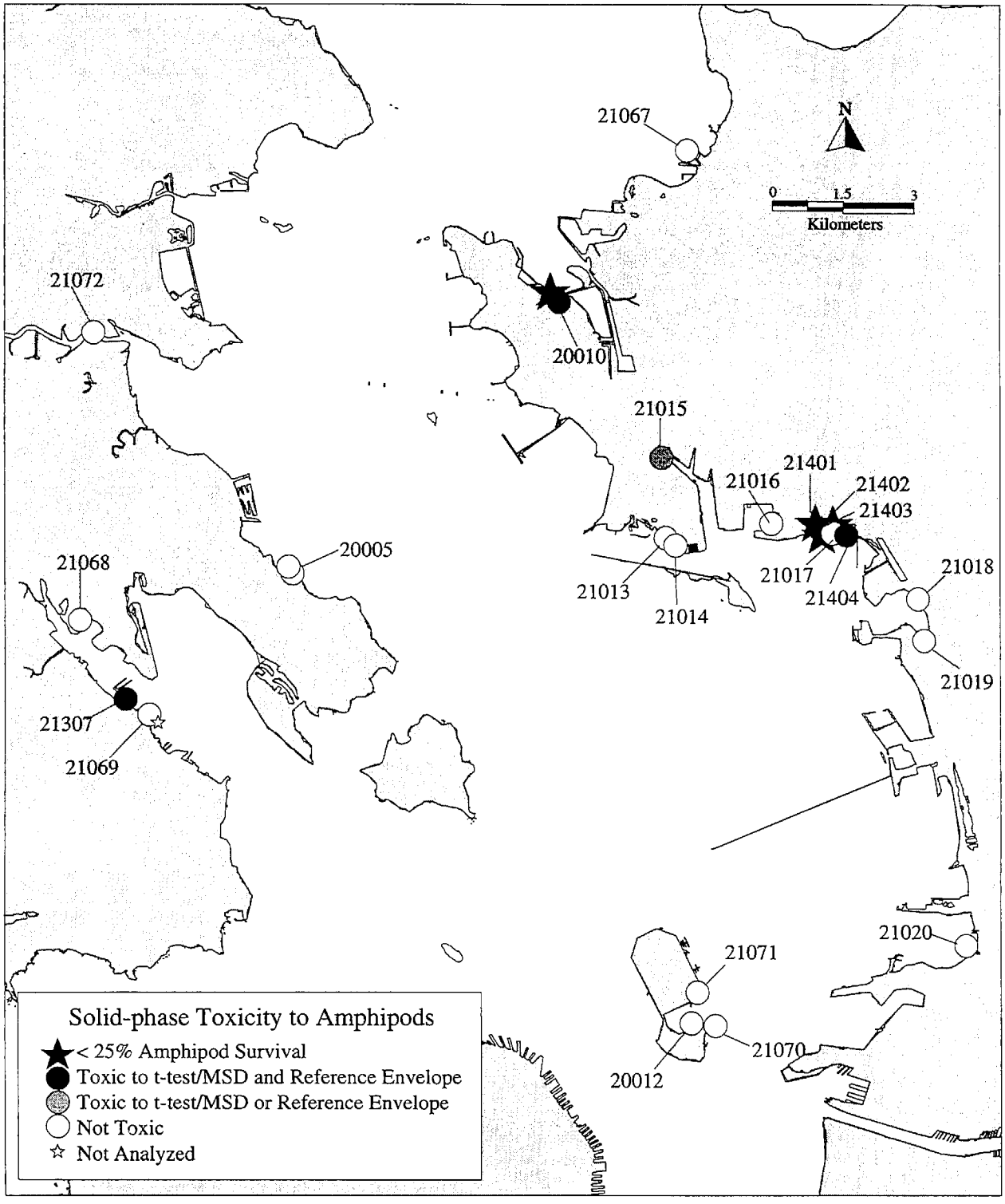

Figure 4c. Results of Amphipod Toxicity Screening for Stations in Central San Francisco Bay.

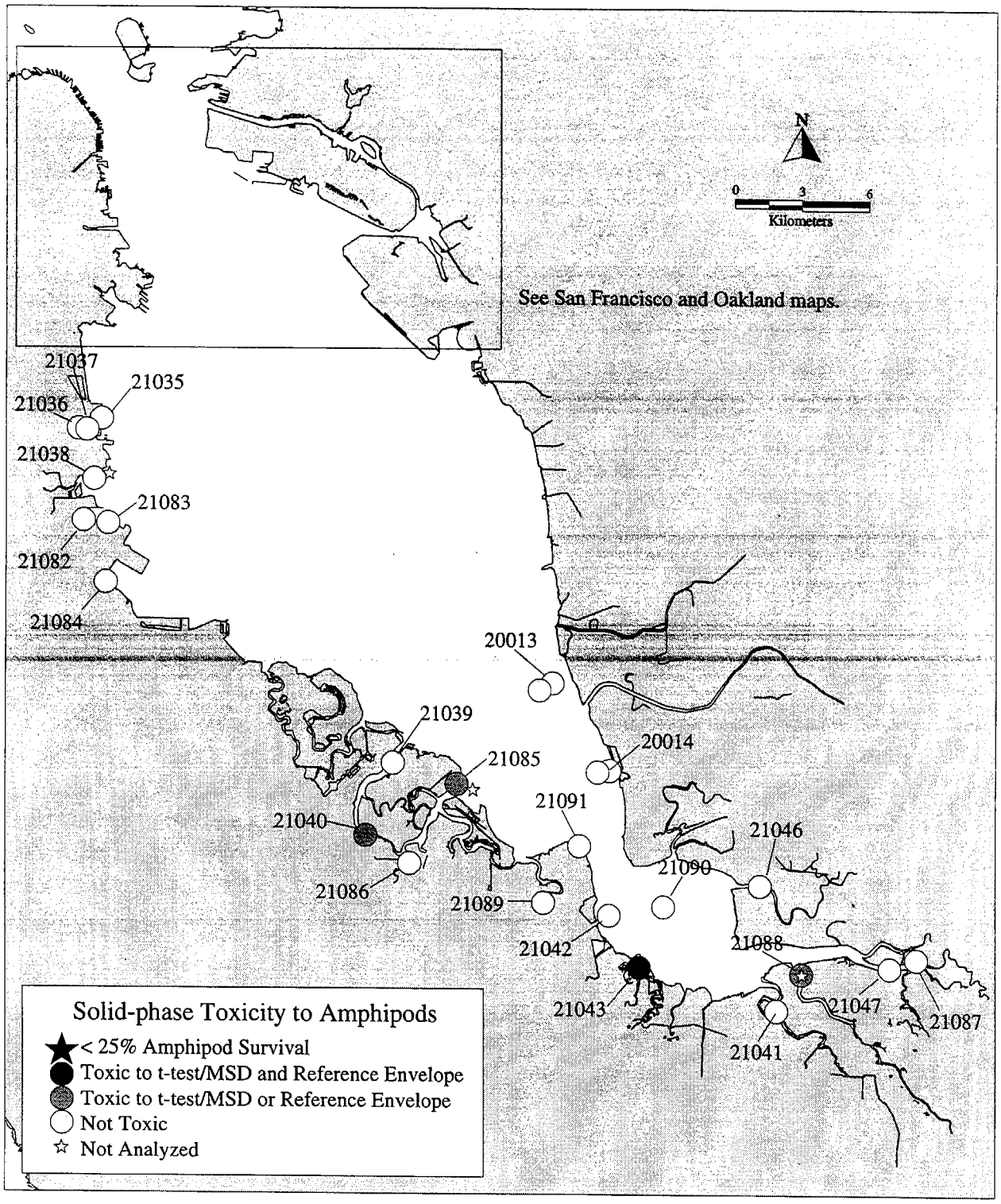

Figure 4d. Results of Amphipod Toxicity Screening for Stations in South San Francisco Bay.

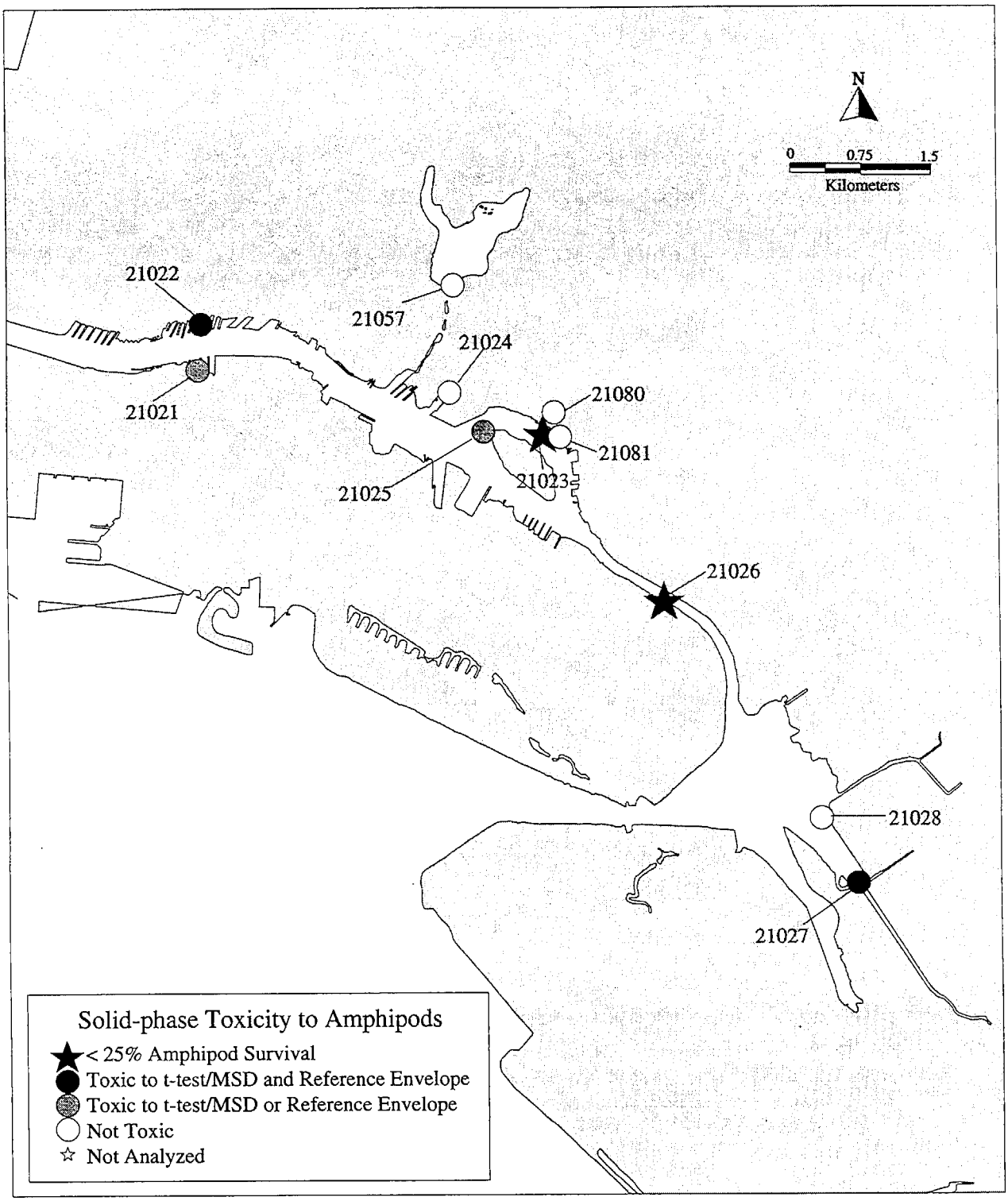

Figure 4e. Results of Amphipod Toxicity Screening for Stations in Oakland.

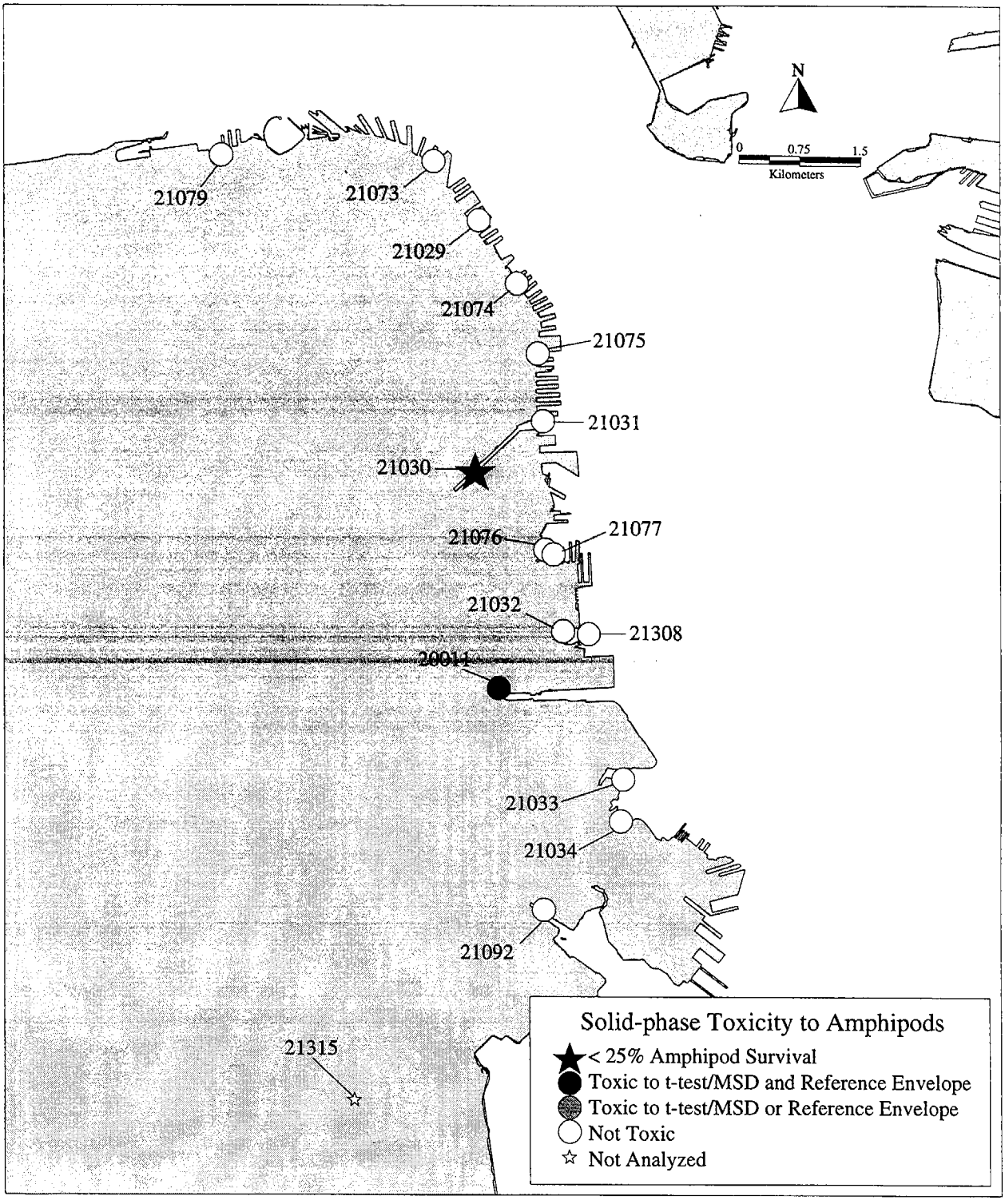

Figure 4f. Results of Amphipod Toxicity Screening for Stations in San Francisco.

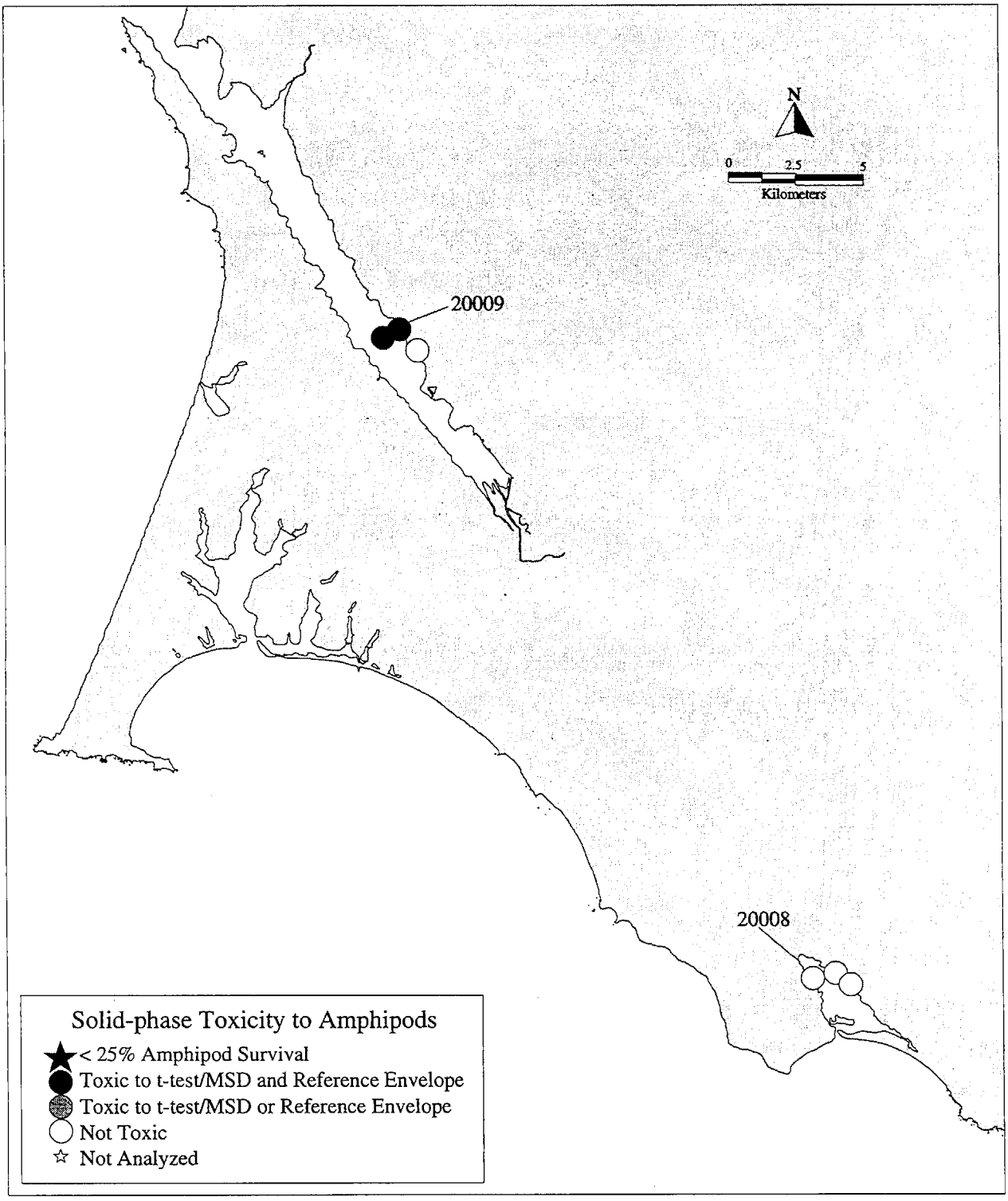

Figure 4g. Results of Amphipod Toxicity Screening for Stations in Tomales Bay and Bolinas Lagoon.

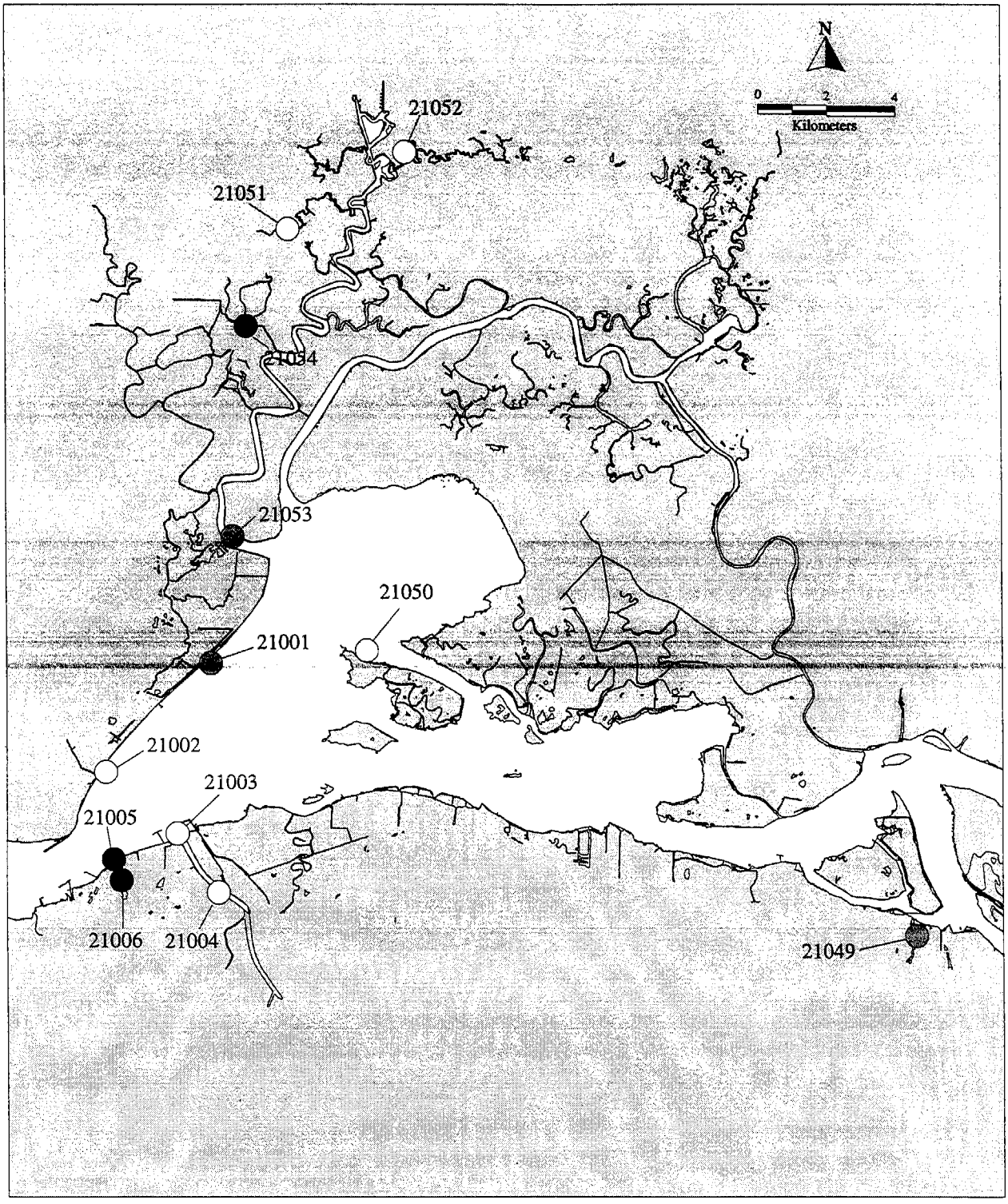

Figure 5a. Results of Porewater Toxicity Screening for Stations in Suisun Bay.

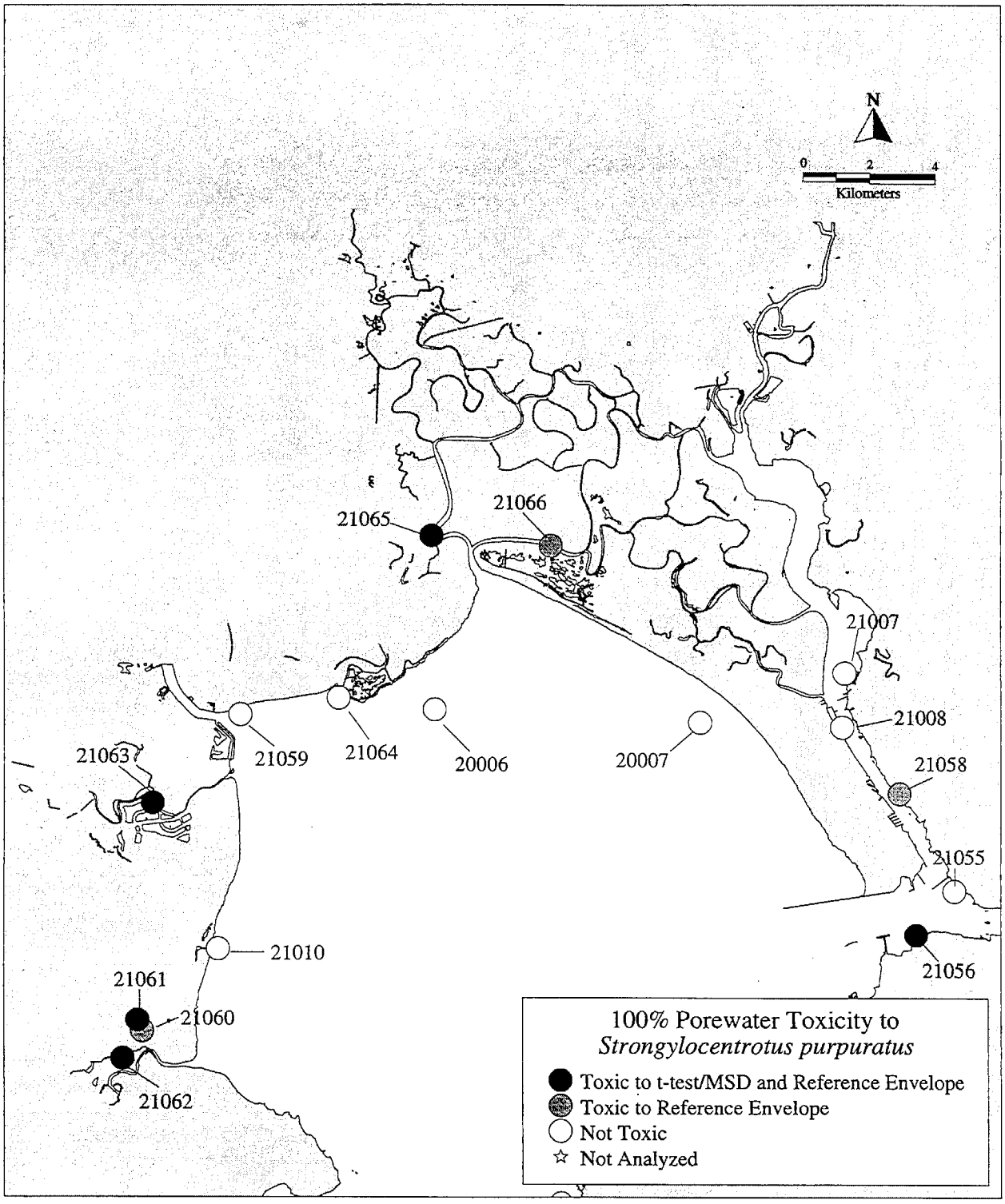

Figure 5b. Results of Porewater Toxicity Screening for Stations in San Pablo Bay.

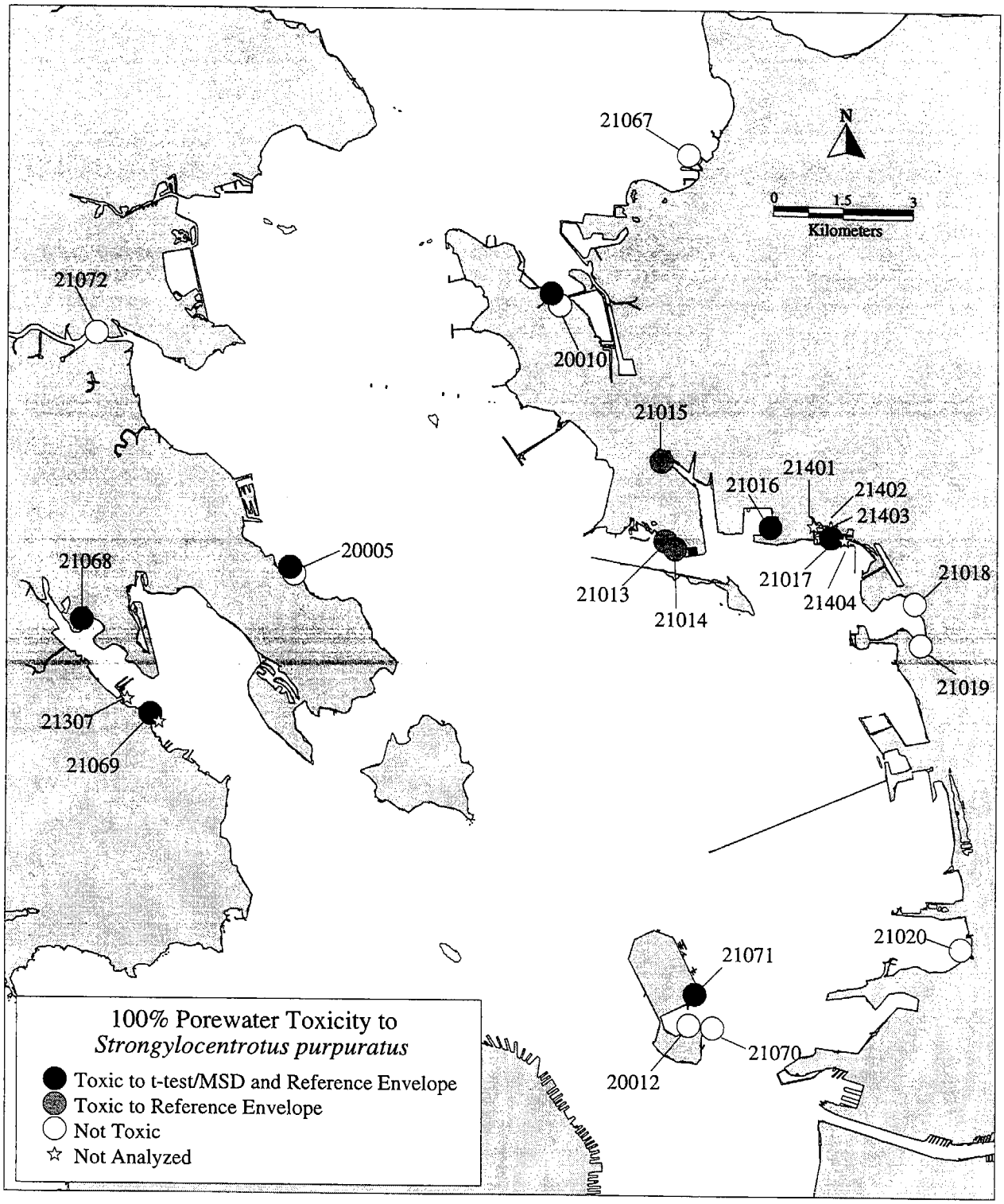

Figure 5c. Results of Porewater Toxicity Screening for Stations in Central San Francisco Bay.

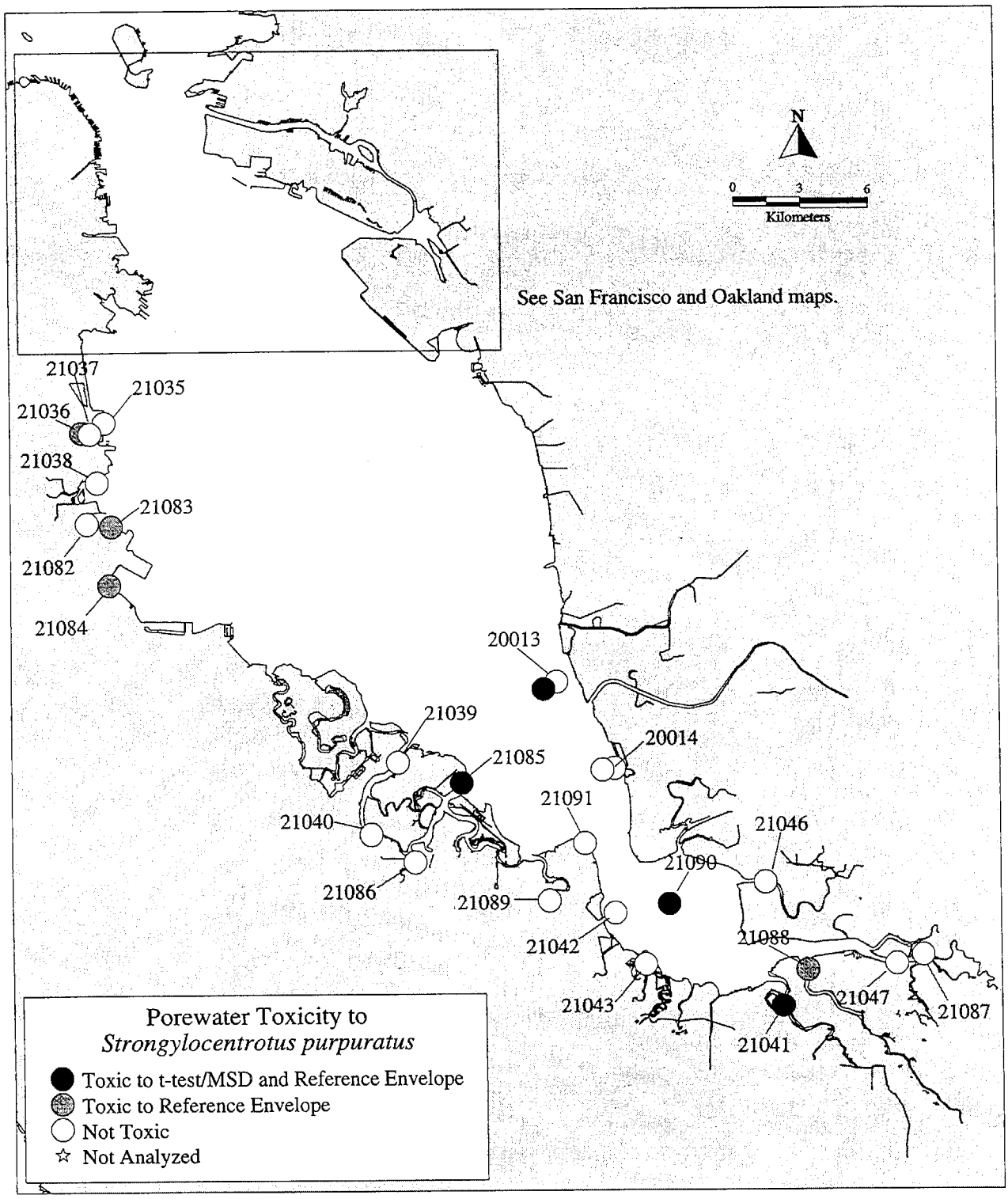

Figure 5d. Results of Porewater Toxicity Screening for Stations in South San Francisco Bay.

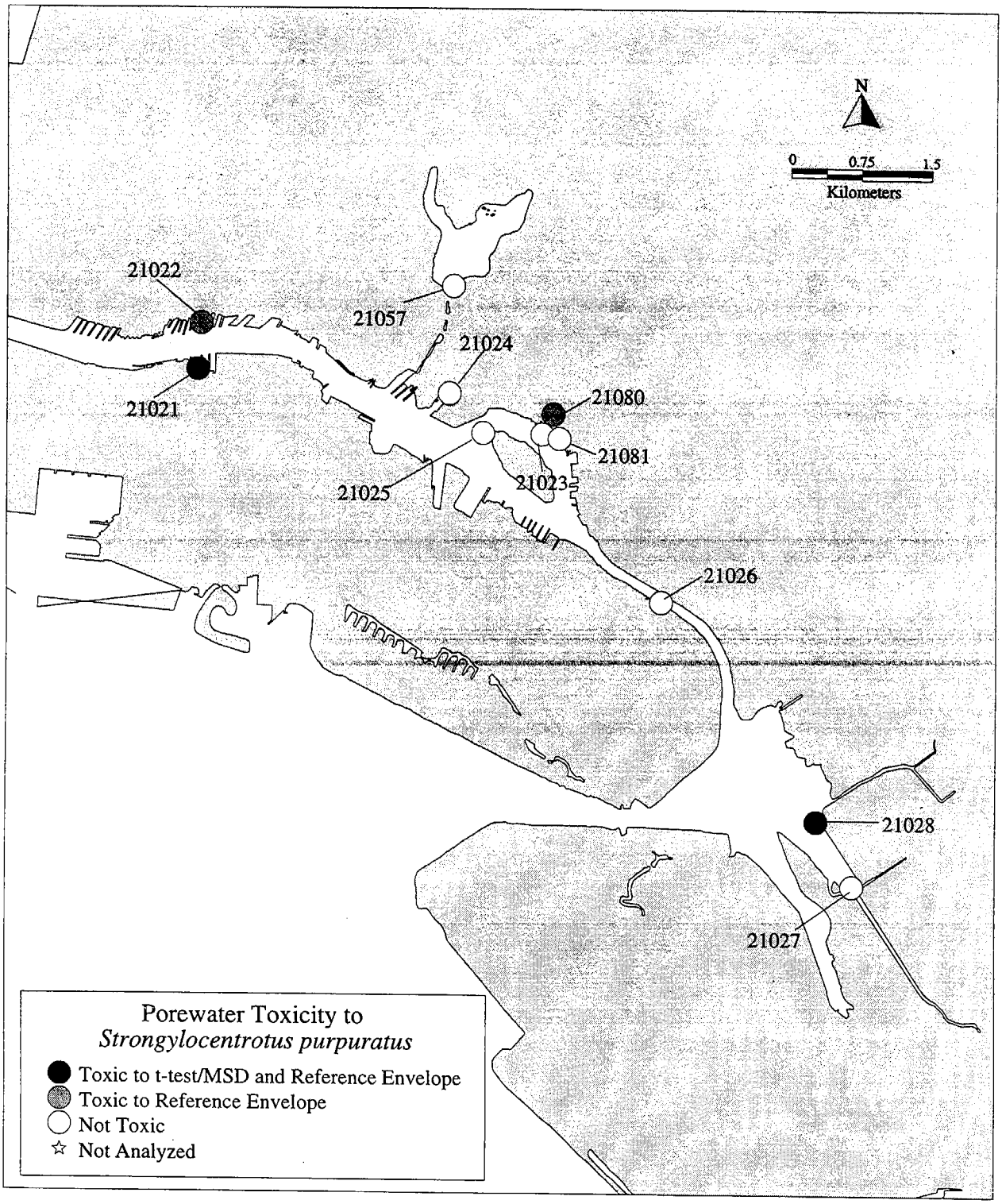

Figure 5e. Results of Porewater Toxicity Screening for Stations in Oakland.

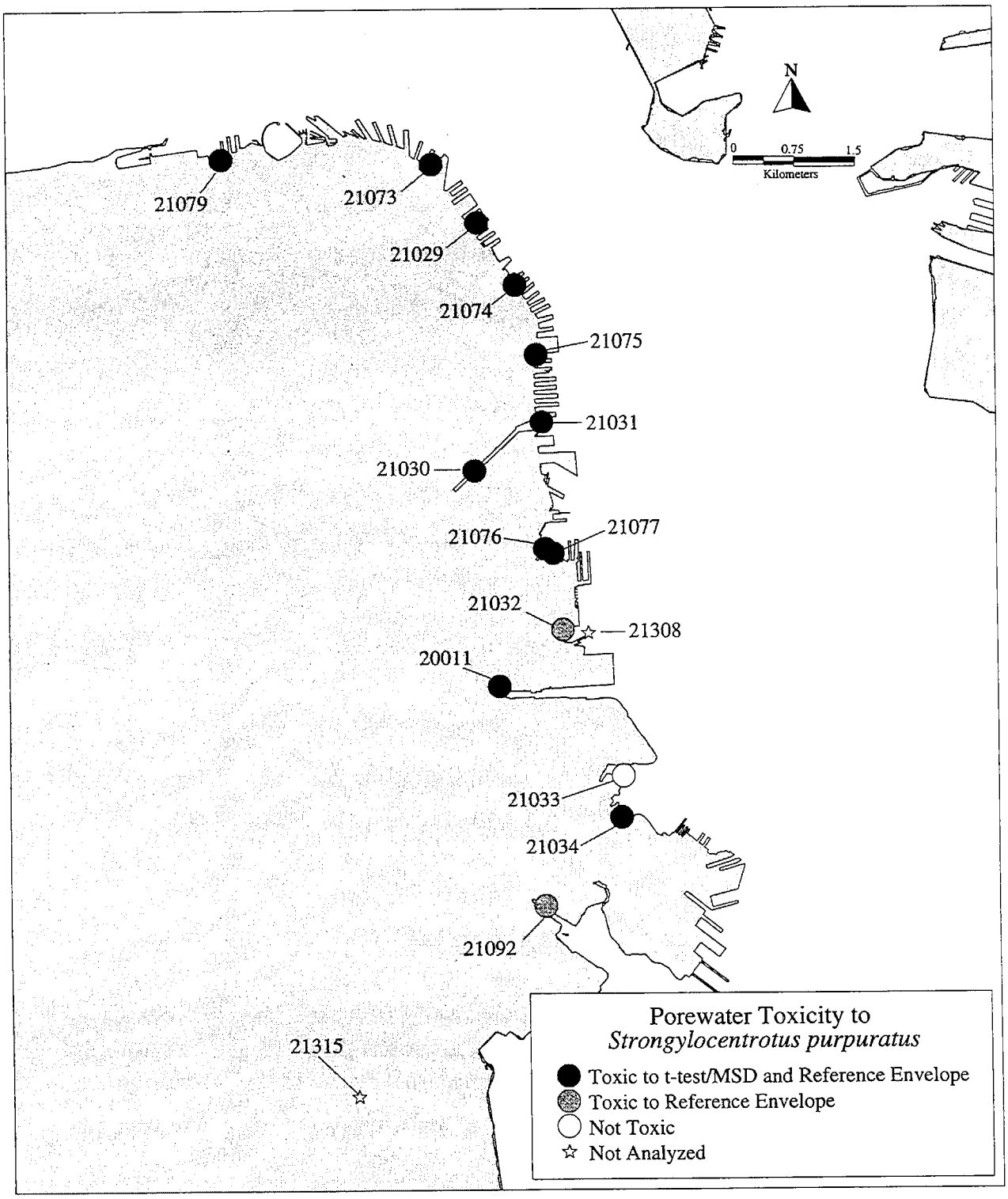

Figure 5f. Results of Porewater Toxicity Screening for Stations in San Francisco.

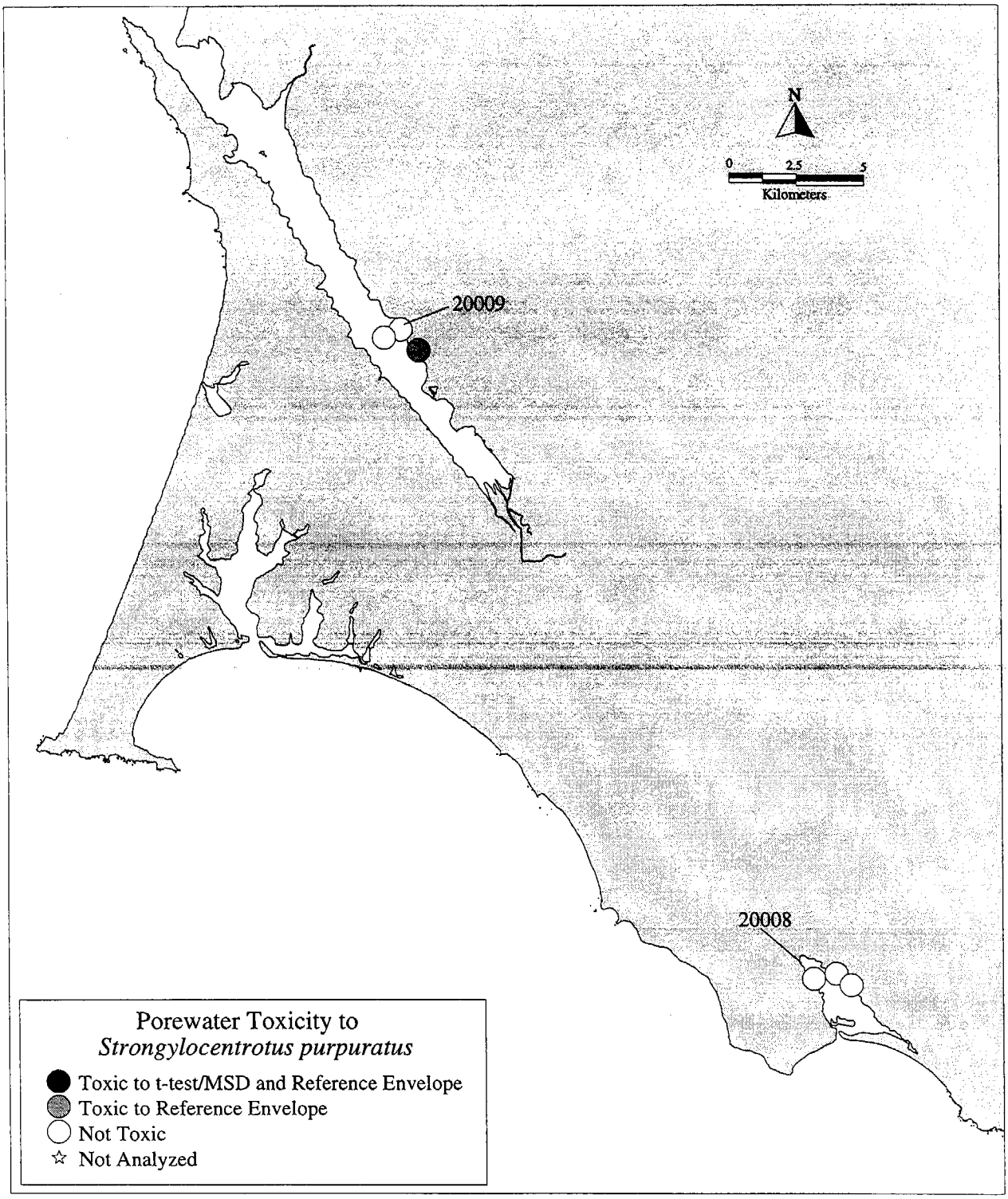

Figure 5g. Results of Porewater Toxicity Screening for Stations in Tomales Bay and Bolinas Lagoon.

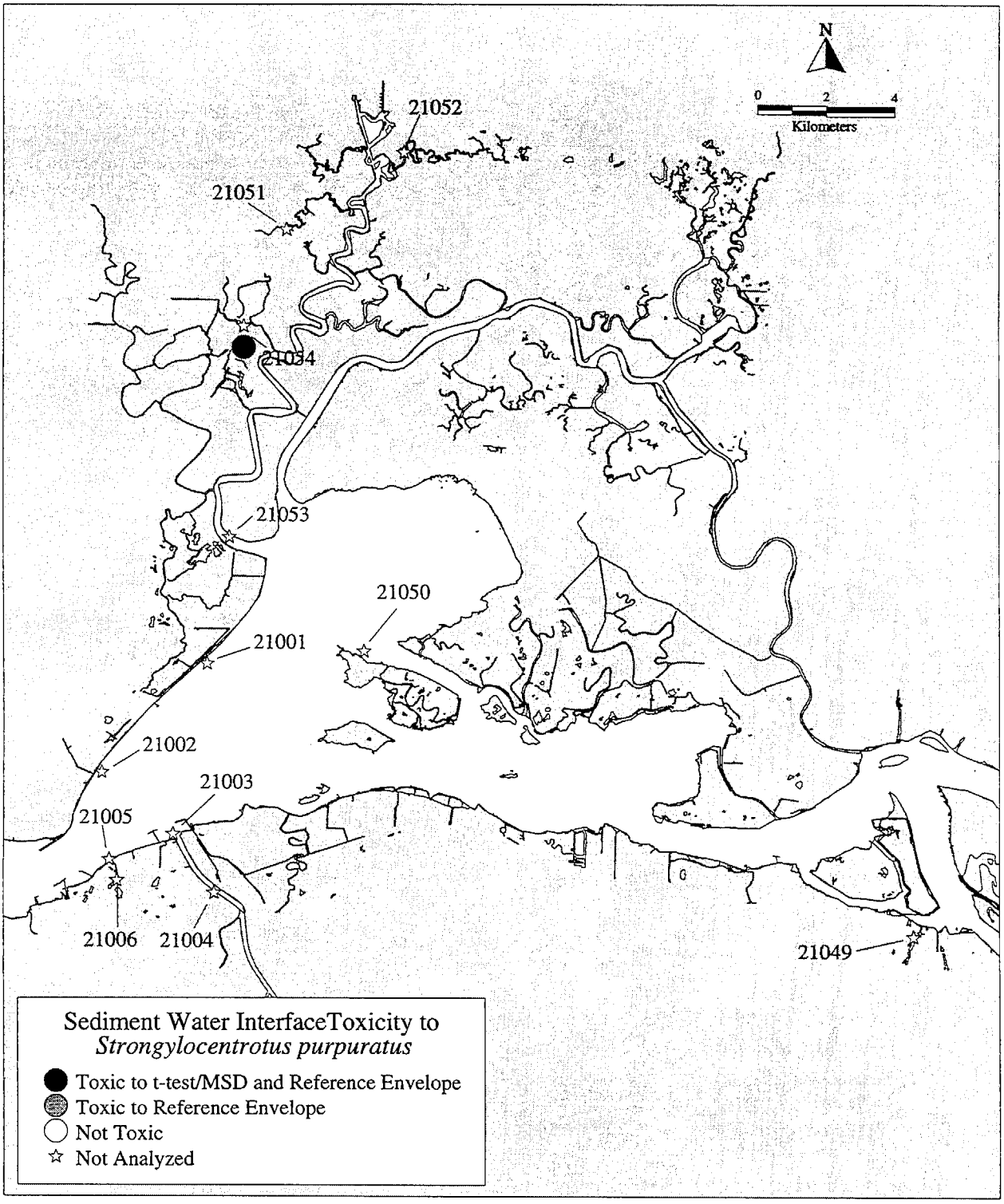

Figure 6a. Results of Sediment-Water Interface Toxicity Screening for Stations in Suisun Bay.

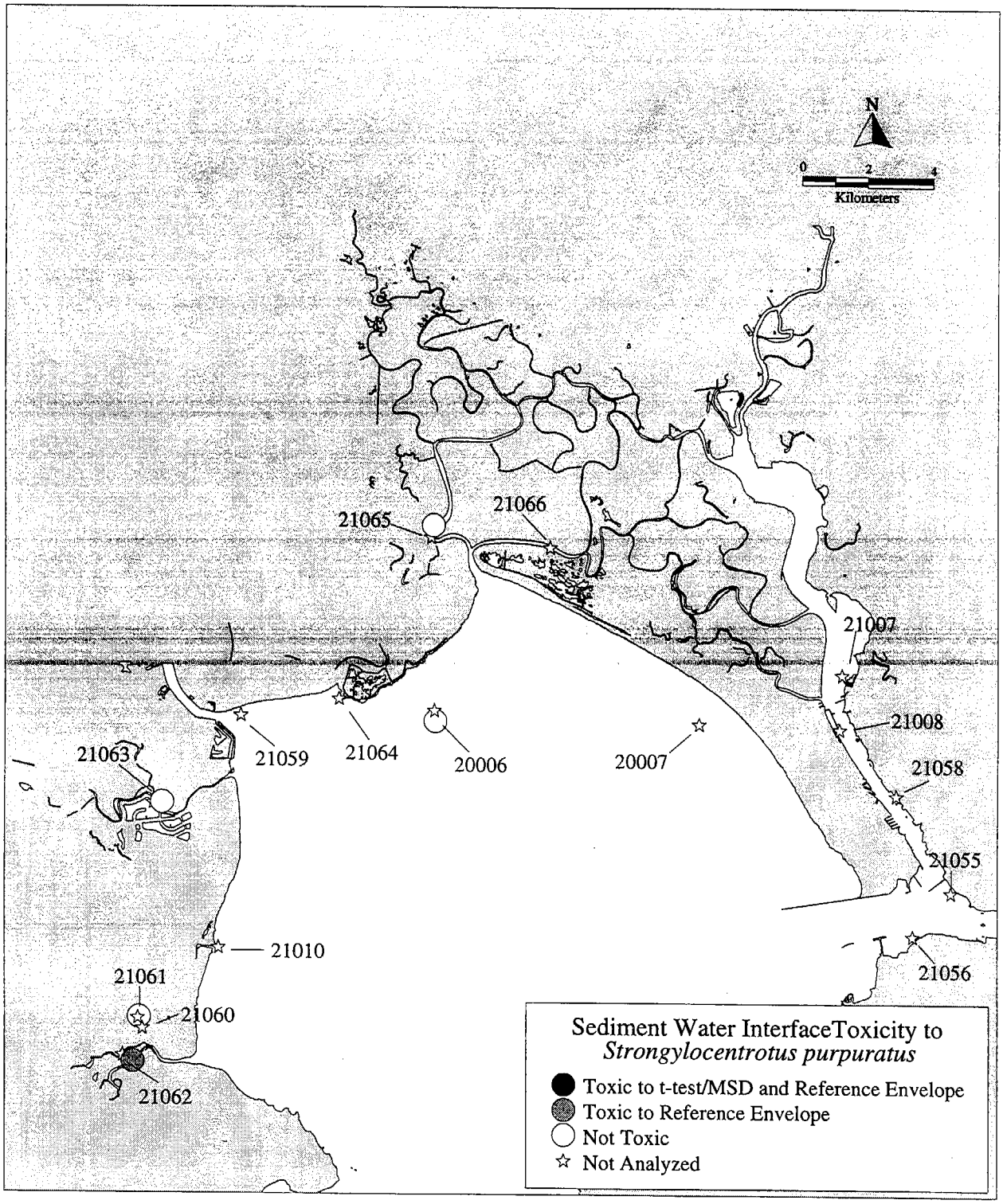

Figure 6b. Results of Sediment-Water Interface Toxicity Screening for Stations in San Pablo Bay.

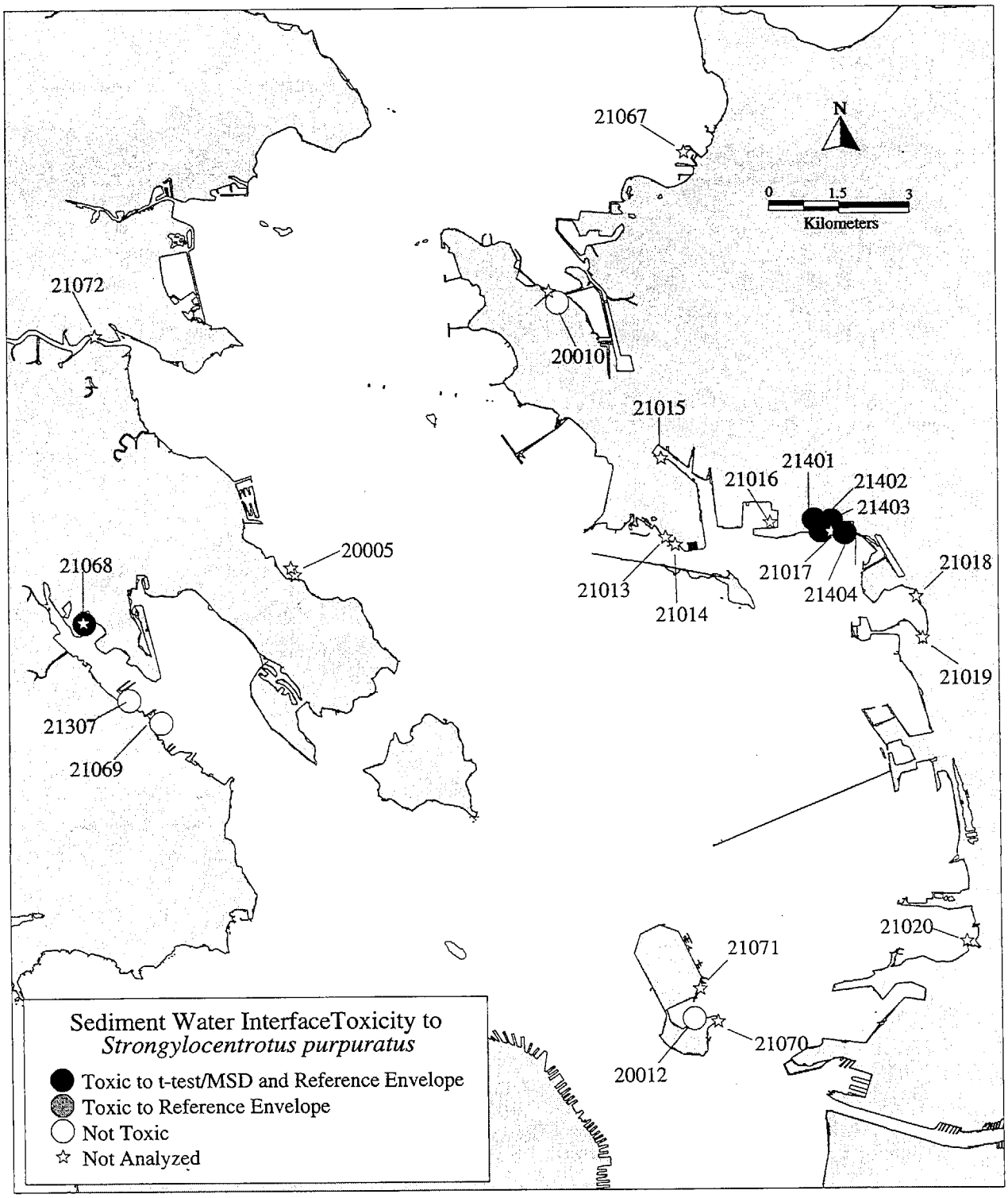

Figure 6c. Results of Sediment-Water Interface Toxicity Screening for Stations in Central San Francisco Bay.

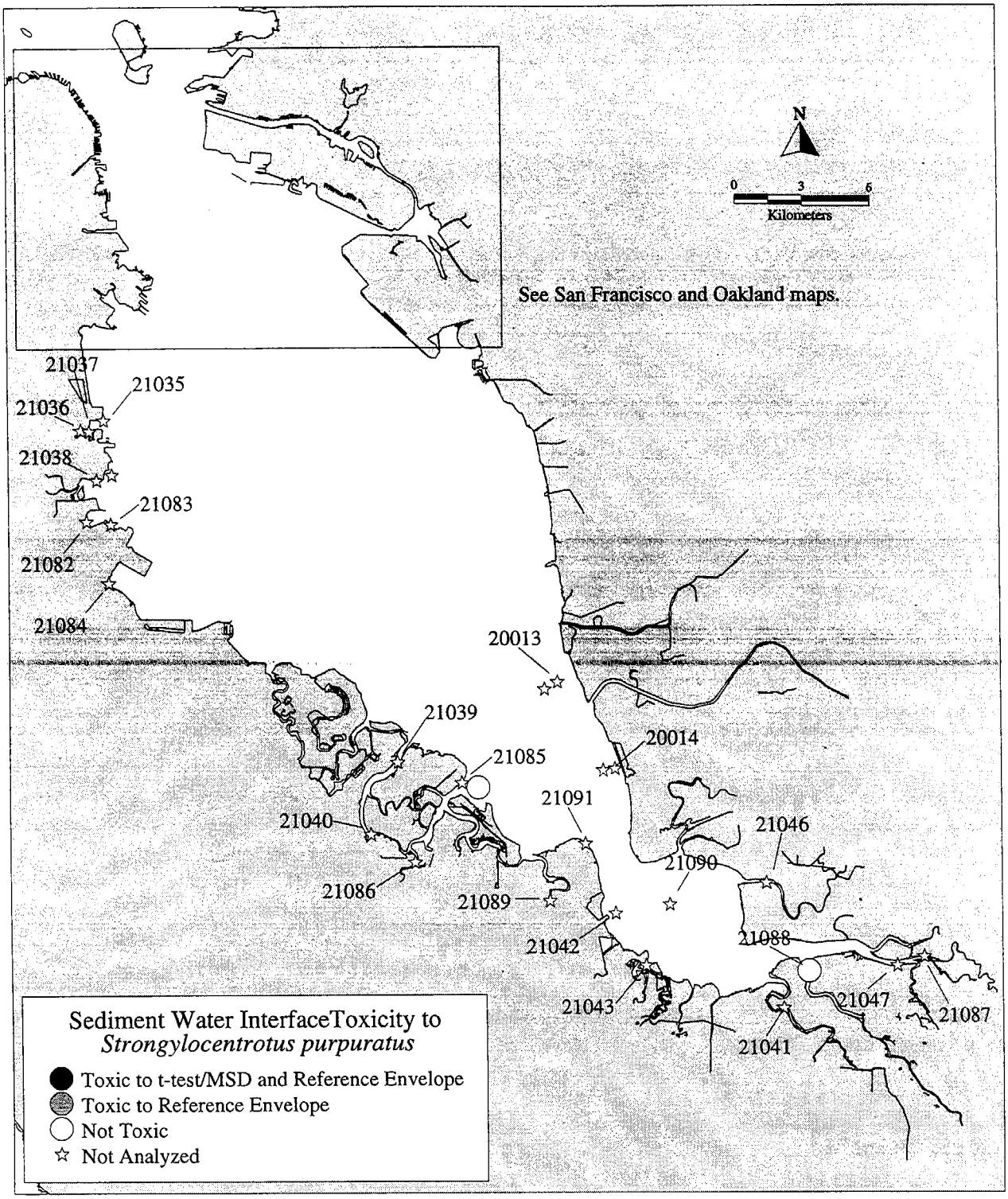

Figure 6d. Results of Sediment-Water Interface Toxicity Screening for Stations in South San Francisco Bay.

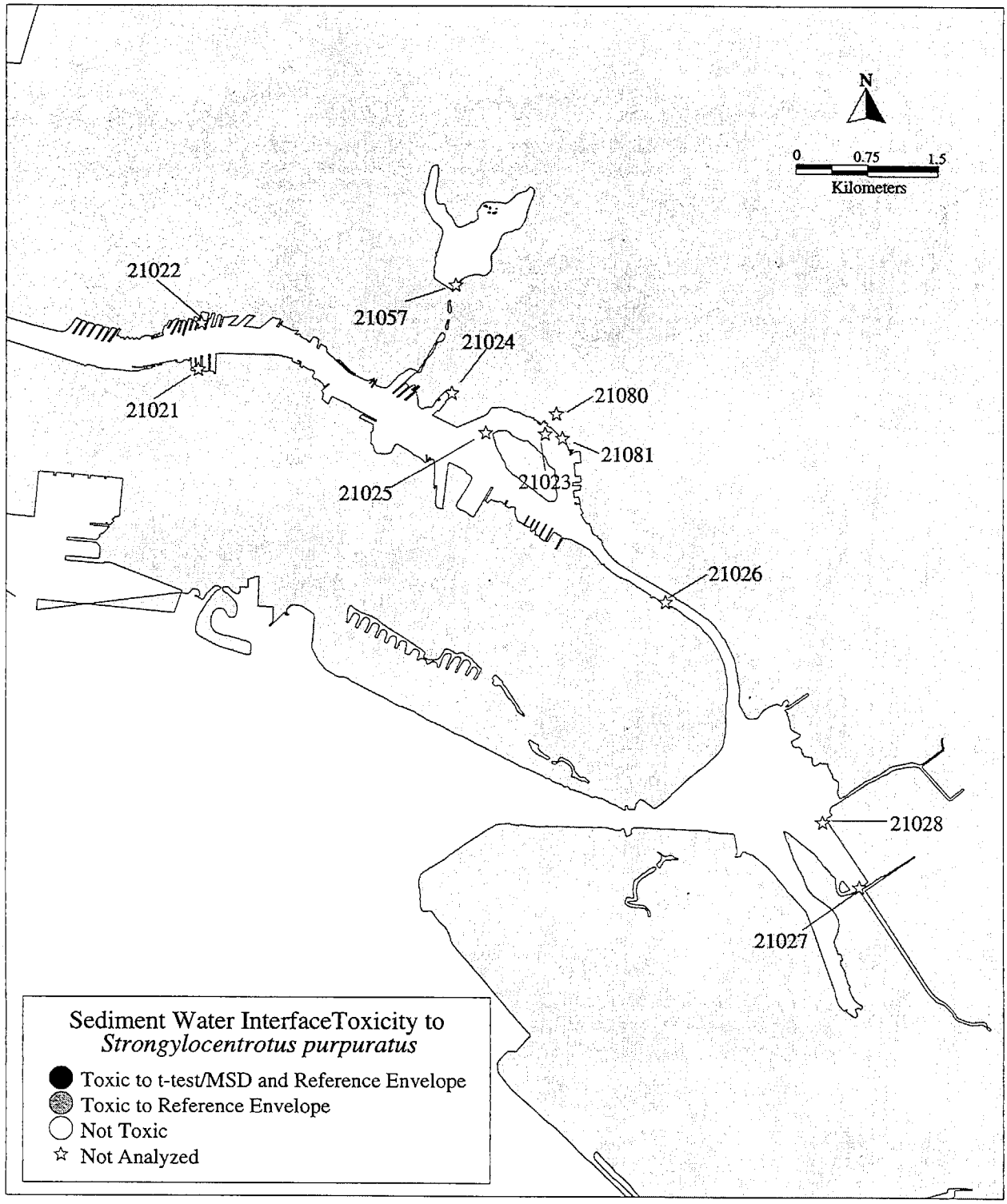

Figure 6e. Results of Sediment-Water Interface Toxicity Screening for Stations in Oakland.

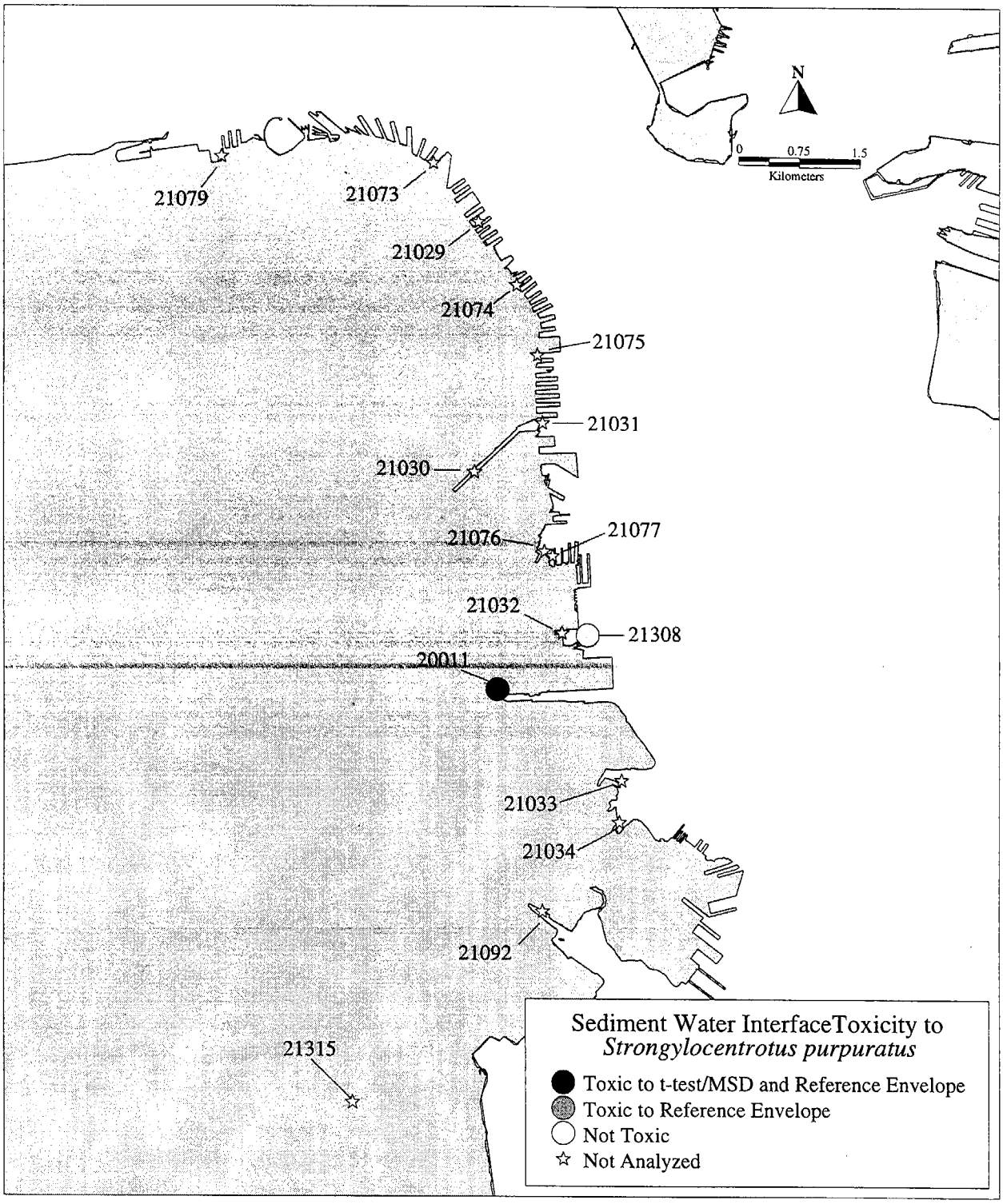

Figure 6f. Results of Sediment-Water Interface Toxicity Screening for Stations in San Francisco.

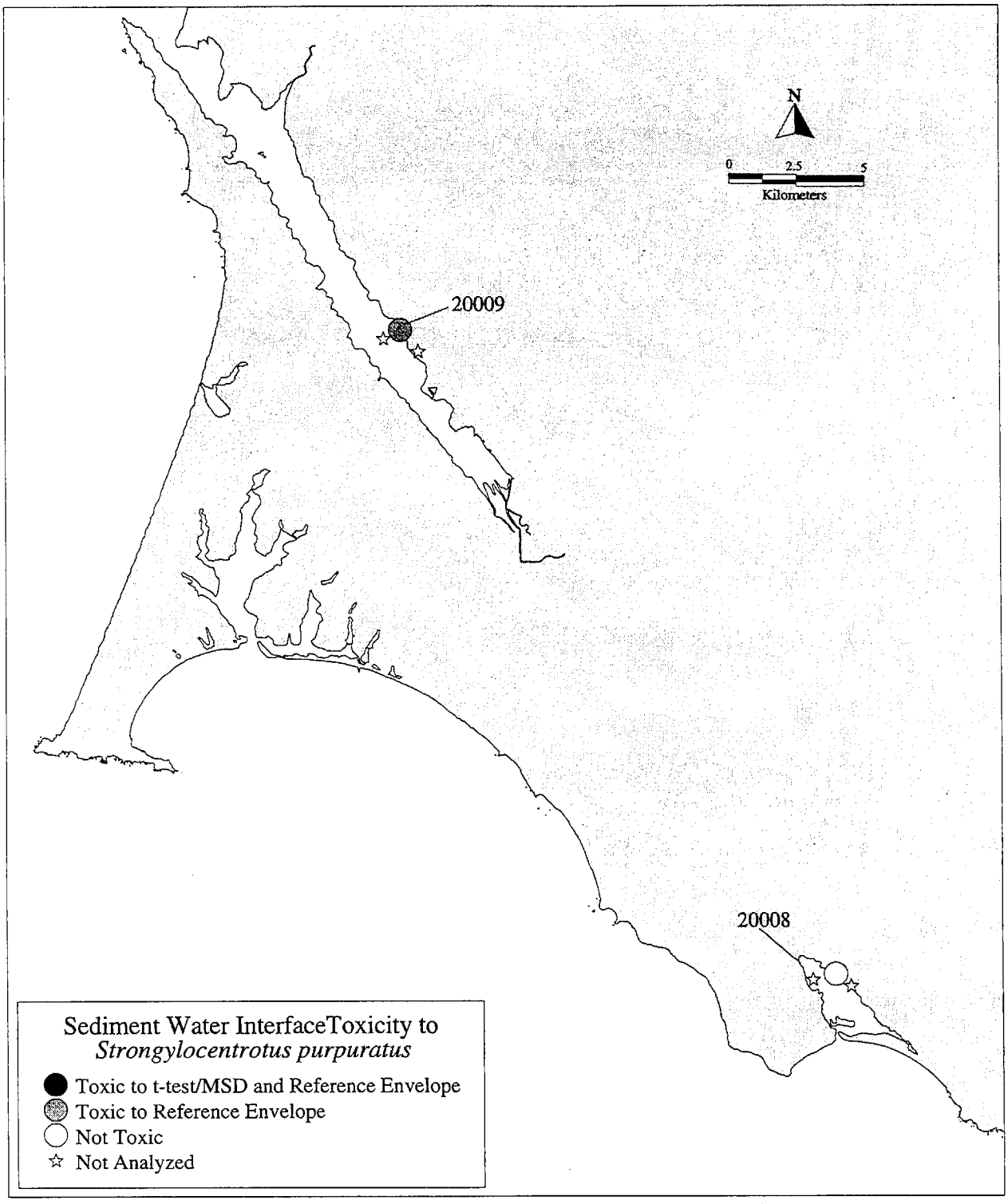

Figure 6g. Results of Sediment-Water Interface Toxicity Screening for Stations in Tomales Bay and Bolinas Lagoon.