Supporting Information

Whitehead et al. 10.1073/pnas.1109545108

SI Methods

The locations (latitude and longitude) of our field sampling sites and dates for sampling at each site are summarized in Dataset S1. Gulf killifish (*Fundulus grandis*) were caught by minnow trap, and tissues were excised immediately. Liver was preserved in RNAlater (Ambion, Inc.) for genome expression (microarray and RNAseq) analysis. Gill tissues were fixed in situ in buffered zinc-based formalin Z-Fix (Anatech LTD).

Satellite Imagery. Satellite imagery was analyzed to provide a coarse but spatially and temporally comprehensive estimation of the timing, location, and duration of coastal oil contamination. Surface oil from the DWH oil spill was detected through the analysis of SAR images, which offer the most effective means of detecting oil remotely. This active radar system operates over large spatial scales in all weather and at all times of day (1, 2). Oil dampens the ocean's smallest capillary waves (3–5 cm in length), yielding black regions in the image attributable to the total lack of microwave backscatter from the sea surface to the sensor, compared with higher backscatter from surrounding regions with waves (Fig. S1). False-positive results are possible from areas with low wind (<3 m/s) and from algal blooms; thus, the use of another satellite sensor or "sea truth" (e.g., wind measurements) is advisable for confirmation of the SAR signal. Only SAR images with distinct signatures, unrelated to these potential artifacts, were used in this study, although even thin oil sheens would potentially yield a dark return because SAR data yield no information about oil thickness. We used SAR measurements from multiple satellites (TerraSARX; ERS-2; CosmoSkymed-1, -2, and -3; Radarsat-1 and -2; Palsar; and Envisat-2). Data were received and processed in real time at the University of Miami Center for Southeastern Tropical Advanced Remote Sensing (CSTARS) laboratory and were further processed at the Louisiana State University Earth Scan Laboratory. The calculated distance from each field sampling site to the nearest oil slick was calculated from the "straight-line" distance from the global positioning system position of the station (Dataset S1) to that of the observed oil across any and all intervening geographical barriers (e.g., Fig. S1). Therefore, calculated distances do not necessarily represent the overall distance oil would have traveled to reach a sample station, although as the calculated distance approaches zero, these two distances (straight line vs. travel distance) become extensionally equivalent.

Analytical Chemistry. Analytical chemistry of water, tissue, and sediment samples was performed to offer detailed characterization of exposure to contaminating oil (data reported in Dataset S2). Sample dates and locations are summarized in Dataset S1. Two liters of water was collected subsurface in 1-L amber-glass jars from each sample site and date, and it was kept at 4 °C until extraction, which was performed within 1 wk of collection. Tissues (whole fish) were collected from each of the field sites from the second (June 2010) and third (August 2010) sampling time points and frozen at −20 °C until extraction. Sediment was collected from each of the field sites after the final sampling time point (September 2010) in 8-oz glass jars and frozen at −20 °C until extraction.

The sediment extraction procedure is as follows. Approximately 30 g of sediment/soil was accurately weighed (to the nearest 0.01 g) into a precleaned 500-mL beaker. The material was homogenized with anhydrous sodium sulfate sample until a "dry" sand-like matrix was created. One milliliter of surrogate standard was spiked into the sample, followed by the addition of 100 mL of pesticide-grade dichloromethane (DCM). The sample mixture was sonicated (60% intensity) for ∼10 min and allowed to settle for 15 min. The solvent was poured over a sodium sulfate funnel to remove any water and drained into 500-mL flat-bottomed flasks. The extraction process was repeated two more times, followed by rinsing the funnel with 25 mL of DCM. The flask was placed on a Buchi evaporative system and reduced to a final volume of 5–10 mL of DCM. The DCM concentrate was pipetted from the flask, placed into a 10-mL microextraction thimble, and reduced to a final volume of 1 mL using a nitrogen blow-down system. The 1-mL extract was transferred to a 2-mL autosampler vial and spiked with 10 μL of internal standard solution. Autosampler vials were stored at 4 °C until ready for analysis.

The water extraction procedure is as follows. Approximately 1,000 mL of water was accurately weighed (to the nearest 1.0 mL) into a precleaned 20,000-mL separatory funnel. One milliliter of surrogate standard was spiked into the sample, followed by the addition of 100 mL of pesticide-grade DCM. The sample mixture was hand-shaken for ∼10 min and allowed to settle for 15 min. The solvent in the bottom of the funnel was drained through a sodium sulfate funnel to remove any water and drained into a 500-mL flat-bottomed flask. The extraction process was repeated two more times, followed by rinsing the funnel with 25 mL of DCM. The flask was placed on a Buchi evaporative system and reduced to a final volume of 5–10 mL of DCM. The DCM concentrate was pipetted from the flask, placed into a 10-mL microextraction thimble, and reduced to a final volume of 1 mL using a nitrogen blow-down system. The 1-mL extract was transferred to a 2-mL autosampler vial and spiked with 10 μL of internal standard solution. Autosampler vials were stored at 4 °C until ready for analysis.

The tissue extraction procedure is as follows. Approximately 5–10 g of tissue was accurately weighed to the nearest 0.01 g into a precleaned 500-mL beaker. The material was homogenized with anhydrous sodium sulfate sample until a dry sand-like matrix was created. One milliliter of surrogate standard was spiked into the sample, followed by the addition of 50 mL of pesticide-grade DCM. The sample mixture was sonicated (60% intensity) for ∼10 min and allowed to settle for 15 min. The solvent was poured over a sodium sulfate funnel to remove any water and drained into a 250-mL flat-bottomed flask. The extraction process was repeated two more times, followed by rinsing the funnel with 25 mL of DCM. The flask was placed on a Buchi evaporative system and reduced to a final volume of 3–5 mL of DCM. The DCM extract was exchanged to hexane with ∼25 mL of pesticide-grade hexane. The flask was returned to the evaporation system and evaporated down to a final volume of 2–5 mL of hexane. The sample was fractionated on an alumina/silica gel column by placing the 2- to 5-mL hexane aliquot on the aluminum/silica gel column, which was then rinsed with high-purity hexane. The flow of hexane was stopped before exposing the silica gel to air. This fraction, which contained alkanes, was collected in a graduated thimble. The alumina/silica gel column was then rinsed with 50% DCM and 50% hexane. The solvents were allowed to elute completely in a separate extraction thimble. This fraction contained the PAHs. The alkane and PAH fractions were combined and concentrated to 1.0 mL under a gentle stream of nitrogen and stored in a 2-mL autosampler vial (4 °C) until GC/MS analysis.

All sample extracts were analyzed using an Agilent 7890A Gas Chromatography system (Agilent Technologies, Inc.) config-

ured with a 5% diphenyl/95% (vol/vol) dimethyl polysiloxane high-resolution capillary column (30 m, 0.25-mm inner diameter, 0.25-μm film) directly interfaced to an Agilent 5975 inert XL MS detector system (Agilent Technologies, Inc.). An Agilent 7638B series Auto Injector (Agilent Technologies, Inc.) was used for sample introduction into the GC/MS system. The GC flow rates were optimized to provide a required degree of separation, which includes near-baseline resolution of n-C17 and pristine, and baseline resolution of n-C18 and phytane. The injection temperature was set at 250 °C, and only high-temperature and low-thermal bleed septa were used in the GC inlet. GC was performed in the temperature program mode with an initial column temperature of 55 °C for 3 min, which was then increased to 280 °C at a rate of 5 °C/min and held for 3 min. The oven was then heated from 280 °C to 300 °C at a rate of 1.5 °C/min and held at 300 °C for 2 min. Total run time was 66.33 min per sample. The interface to the MS was maintained at 280 °C. Ultra-high-purity helium was the carry gas for the GC/MS system.

Spectral data were processed by Chemstation Software (Agilent Technologies, Inc.). Analyte concentrations were calculated based on the internal standard method. Therefore, an internal standard mixture composed of naphthalene-d8, acenaphthene-d10, chrysene-dl2, and perylene-dl2 (usually at a concentration of 10 ng/μL) was spiked into the sample extracts just before analysis. The concentration of specific target oil analytes was determined by a five-point calibration and internal standard method. Standards containing parent (nonalkylated) hydrocarbons were used in the calibration curve. Alkylated homologs were quantified using the response factor of the parent, and were therefore semiquantitative. This was the standard procedure, because alkylated standards were not available.

Genome Expression: Microarrays. Genome expression across sites and time was characterized using custom oligonucleotide microarrays. Genome expression was measured in liver tissues from five replicate individual male fish per site-time treatment (5 biological replicates). Male fish were chosen for genome expression analysis because sampling was conducted during spawning season, when female reproductive condition (and associated liver genome expression) can be highly variable. Microarray probes (60-mer) were designed from contigs constructed from *F. heteroclitus*-expressed sequence tags. *F. heteroclitus* is the Atlantic coast-distributed sister species of Gulf coast-distributed *F. grandis* (3). Microarrays included probes for 6,800 unique EST sequences, each printed in duplicate on 15,000 element custom Agilent microarrays (design ID no. 027999) (Agilent Technologies, Inc.). Total RNA was extracted using TRIzol reagent, antisense RNA (aRNA) prepared using the amino allyl aRNA amplification kit (Ambion, Inc.), and purified aRNA coupled to Alexa Fluor dyes (Alexa Fluor 555 and 647; Molecular Probes, Inc.), and it was hybridized to custom microarrays for 18 h at 60 °C in a balanced loop design. Microarray images were captured using a Packard Bioscience ScanArray Express (PerkinElmer, Inc.) microarray scanner, and images were processed using Imagene (Biodiscovery, Inc.). Spots that were too bright (saturated) or too faint (below 2 SDs above background intensity) were excluded from normalization, resulting in a final set of 3,296 probes included for normalization and statistical analysis (Dataset S3). Data were lowess-normalized and then mixed model-normalized using linear mixed models to account for fixed (dye) effects and random (array) effects. Normalized data were then analyzed using mixed model ANOVA, with "site" [Grand Terre (GT), Bay St. Louis (BSL), Belle Fontaine Point (BFP), Bayou La Batre (BLB), Mobile Bay (MB), and Fort Morgan (FMA)] and "sampling time" (sampling trips 1, 2, and 3) (Dataset S1) as main effects, including an interaction (site-by-time) term. "Dye" was considered a fixed effect, and "array" and "replicate individual within site-time treatment" ($n = 5$) were treated as random effects. The false discovery rate was estimated using Q-value (4). Principal components analysis was performed using MeV (5). GO enrichment was tested using DAVID (6).

Genome Expression: RNAseq. Transcript abundance was compared between liver mRNA from three replicate fish (RNA was not pooled) from the GT site from June 28, 2010, and mRNA from two control samples. The two control samples are composed of pooled liver mRNA from six and eight individuals, respectively, collected in April 2008. The individuals for one control sample were collected (2 each) from three sites west of the Mississippi river, including Port Aransas, Texas; Cocodrie, Louisiana; and LeeVille, Louisiana. The individuals for the second control sample were collected (2 each) from four sites west of the Mississippi River, including Dauphin Island, Alabama; Weeks Bay, Alabama; Santa Rosa Island, Florida; and St. Teresa, Florida. All RNA samples were sequenced on the Illumina Gene Analyzer platform (Expression Analysis, Inc.), and the resulting short-read data were summarized in fastq format. Short reads with more than two uncalled bases were removed. Each read was cut whenever a position fell below a minimum quality score of 10 or if the average of the qualities of a position and its two neighbors fell below 20, and the largest remaining fragment was used.

Quantitative transcript abundance analysis was initiated by mapping filtered short reads to target sequences (6,810 unique *F. heteroclitus* target EST sequences, Dataset S5) using the Bowtie short read alignment software (7). A custom Perl script determined the number of fragments mapped to each target sequence. The Bioconductor package DESeq (version 2.8) (8) was then used to determine statistical significance of each differentially expressed target using a negative binomial method with *P* values adjusted by the Benjamini–Hochberg procedure. The three GT site samples were identified as a single "Exposed" class to DESeq, and the two pooled samples were identified as a single "Control" class.

Gill Morphology and Protein Expression: Field Study. Male and female fish were sampled from all field sites for analysis of CYP1A protein expression in the gills. Tissues were fixed immediately in Z-Fix, stored on ice, and held at room temperature before further processing. Gill tissues from at least three fish per site per sampling time were dehydrated in ascending grades of histology-grade ethanol. Tissues were then transferred to a t-butanol bath before clearing in Histochoice Clearing Agent (Amersco) and embedding in Paraplast (Sigma). Tissues were cut along the longitudinal axis at a thickness of 4 μm using an American Optical 820 microtome and transferred onto poly-L-lysine–coated microscope slides. After rehydration, tissues were processed for antigen retrieval by microwave in Tris-buffered saline (pH 9.0) and blocked. Tissues were then probed with mAb C10-7 against fish CYP1A (9). Sections were counterprobed using the Vectastain ABC immunoperoxidase system (Vector Laboratories), utilizing the ImmPACT Nova RED peroxidase substrate kit (Vector Laboratories) to visualize the CYP1A protein in red. Tissue sections were counterstained with Vector Hematoxylin QS (Vector Laboratories). Slides were then observed with a Leica DM RXA2 microscope (Leica Microsystems), and images were captured with a Spot Insight 4 megapixel camera (Diagnostic Instruments). Representative images were captured at a magnification of 40×.

Early Life-Stage Experiments. Approximately 20 L of water was collected (in coordination with collection of water for analytical chemistry; Dataset S2) subsurface from field sites on the dates indicated in Dataset S1. Water was stored in airtight stainless-steel soda kegs and kept at 4 °C until experiments were conducted. Water samples from GT and BLB were utilized in laboratory exposures of *F. grandis* embryos obtained by in vitro

fertilization using ova and spermatozoa collected from a brood stock of unexposed adult *F. grandis* derived from Cocodrie, Lousiana before oiling. Cocodrie parental stock fish were maintained at Louisiana State University, where they were held in the aquatics facility at the Department of Biological Sciences in 400-L tanks maintained at 17 parts per thousand (ppt) water (Instant Ocean) under recirculating conditions.

Following fertilization, 20 embryos were randomly transferred in triplicate to one of the six field-collected waters (2 field sites × 3 time points) at 3 h postfertilization. Embryos were also exposed to a laboratory control consisting of artificial 17 ppt water. Larvae at 24 d postfertilization were sampled and fixed in Z-Fix solution. After fixation, tissues were prepared, sectioned, and stained with the mAb C10-7, as described in the previous section.

1. Brekke C, Solberg AHS (2005) Oil spill detection by satellite remote sensing. *Remote Sensing of Environment* 95:1–13.
2. Fingas MF, Brown CE (1997) Review of oil spill remote sensing. *Spill Science and Technology Bulletin* 4:199–208.
3. Whitehead A (2010) The evolutionary radiation of diverse osmotolerant physiologies in killifish (*Fundulus* sp.). *Evolution* 64:2070–2085.
4. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100:9440–9445.
5. Saeed AI, et al. (2006) TM4 microarray software suite. *Methods Enzymol* 411:134–193.
6. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.
7. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.
8. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11:R106.
9. Rice CD, Schlenk D, Ainsworth J, Goksoyr A (1998) Cross-reactivity of monoclonal antibodies against peptide 277-294 of rainbow trout CYP1A1 with hepatic CYP1A among fish. *Mar Environ Res* 46:87–91.

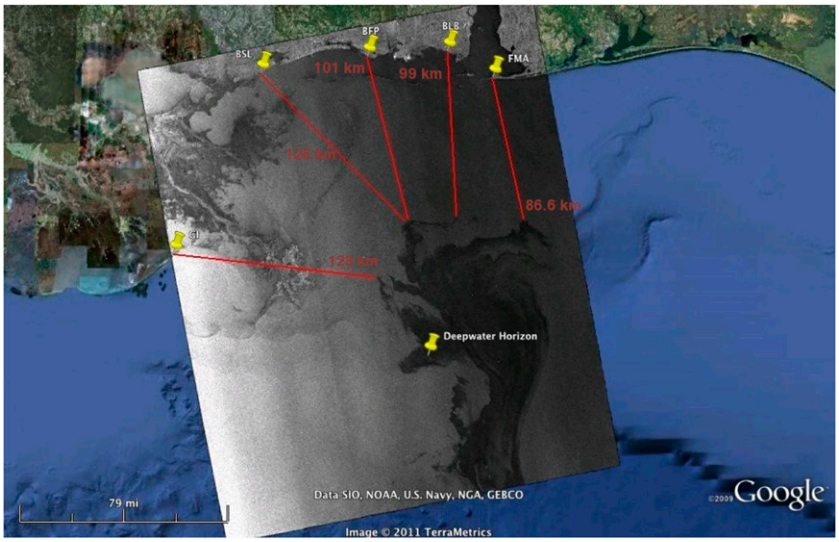

Fig. S1. Representative measurements of the distance from field sites to ocean surface oil according to the CosmoSkymed2 SAR image captured May 13, 2010, at 11:56 UTC (Coordinated Universal Time). Field sites include Grand Terre (GT), Bay St. Louis (BSL), Belle Fontaine Point (BFP), Bayou La Batre (BLB), and Fort Morgan (FMA).

Fig. S2. Oil contaminating the marsh at the GT field site on June 16, 2010 (photograph by B.D.).

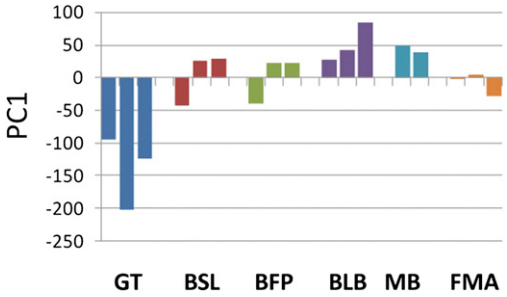

Fig. S3. Expression divergence along principal component 1 (PC1) across consecutive sampling times for the subset of 380 genes that was dose-responsive to PCB exposure in a study by Whitehead et al. (1). Field sites include Grand Terre (GT), Bay St. Louis (BSL), Belle Fontaine Point (BFP), Bayou La Batre (BLB), and Fort Morgan (FMA).

1. Whitehead A, Pilcher W, Champlin D, Nacci D (2011) Common mechanism underlies repeated evolution of extreme pollution tolerance. *Proc R Soc B*, 10.1098/rspb.2011.0847.

Dataset S1. Sites, precise locations, and sampling dates for three field sampling trips

Dataset S1

Dataset S2. Analytical chemistry of subsurface water samples, tissue samples (whole fish), and sediment samples

Dataset S2

Dataset S3. Genome expression microarray data: All probes included in the analysis, including the target EST sequence, probe sequence, annotation, average expression within each treatment (average of $n = 5$ replicate samples within each site-by-time treatment), and results from statistical analyses

Dataset S3

Dataset S4. Results of GO enrichment analysis using DAVID for the subset of genes that were divergently expressed at the GT site coincident with oil contamination

Dataset S4

Dataset S5. Genome expression RNAseq data: All gene targets included in the analysis, including the target EST sequence, annotation, fold difference in transcript abundance between the average of three replicate fish from GT sample time 2 (June 28, 2010) and two replicate reference RNA pools, and adjusted *P* values

Dataset S5