Appendix 1.— Field codes, data scale, and description of species identifiers, traits, and other information compiled in the FishTraits database.

Field Code	Data Scale	Description and Notes
Species Identifiers		
SID	N/A	A unique identification for the species. We concatenated the first letter of the generic name and the first seven letters of the specific name or the first two letters of the generic name and the first six letters of the specific name. A blank SID means the species was not found in the current American Fisheries Society (AFS) list of species (Nelson et al. 2004).
ALTSID	N/A	Alternate species identification. This is essentially the same as SID. However, some species identifications were created from scientific names or other names used in the literature prior to our final checking of species names with the AFS list (Nelson et al. 2004). This identification is also unique and available for all species.
FID	N/A	Family identification formed from the first eight letters of the family name (Nelson et al. 2004).
GENUS	N/A	Genus name recognized by AFS (Nelson et al. 2004).
SPECIES	N/A	Species name recognized by AFS (Nelson et al. 2004).
GID	N/A	Genus identification formed from the first eight letters of the genus name.
COMMONNAME	N/A	Common name recognized by AFS (Nelson et al. 2004).
OTHERNAMES	N/A	Other common names encountered in accounts of the species.
NATIVE	Binary	Native status, as determined from the online database of non-indigenous species (Nico and Fuller 1999).
NOTES	Categorical	A unique species number for footnotes and annotations.
ITISTSN	Categorical	Unique Taxonomic Serial Number (TSN) based on the Integrated Taxonomic Information System http://www.itis.gov .
FAMILYNUMBER	Ordinal	This number is unique to families and serves more than as an identifier. It also indicates phylogenetic position of the family relative to other families (Nelson 2006).

		Smaller numbers are associated with more evolutionarily primitive fishes. Because of phylogenetic constraints on traits data, it is often necessary to explicitly recognize phylogeny in traits-based analyses. These numbers are useful under such circumstances.
Trophic Ecology Traits		
NONFEED	Binary	Adults do not feed.
BENTHIC	Binary	Benthic feeder.
SURWCOL	Binary	Surface or water column feeder.
ALGPHYTO	Binary	Algae or phytoplankton, including filamentous algae.
MACVASCU	Binary	Any part of macrophytes and vascular plants.
DETRITUS	Binary	Detritus or unidentifiable vegetative matter.
INVLVFSH	Binary	Aquatic and terrestrial invertebrates including zooplankton, insects, microcrustaceans, annelids, mollusks, etc. This group also includes larval fishes.
FSHCRCRB	Binary	Larger fishes, crayfishes, crabs, frogs, etc.
BLOOD	Binary	For parasitic lampreys that feed mainly on blood.
EGGS	Binary	Eggs of fishes, frogs, etc.
OTHER	Binary	Other diet components distinct from the preceding classes.
Body Size, Reproductive Ecology, and Life History Traits		
A_1_1	Binary	Nonguarders; Open substratum spawners; Pelagophils.
A_1_2	Binary	Nonguarders; Open substratum spawners; Litho-pelagophils.
A_1_3A	Binary	Nonguarders; Open substratum spawners; Lithophils (rock-gravel).
A_1_3B	Binary	Nonguarders; Open substratum spawners; Lithophils (gravel-sand).
A_1_3C	Binary	Nonguarders; Open substratum spawners; Lithophils (silt-mud).
A_1_4	Binary	Nonguarders; Open substratum spawners; Phyto-lithophils.

A_1_5	Binary	Nonguarders; Open substratum spawners; Phytophils.
A_1_6	Binary	Nonguarders; Open substratum spawners; Psammophils.
A_2_3A	Binary	Nonguarders; Brood hiders; Lithophils (rock-gravel).
A_2_3B	Binary	Nonguarders; Brood hiders; Lithophils (gravel-sand).
A_2_3C	Binary	Nonguarders; Brood hiders; Lithophils (mud).
A_2_4A	Binary	Nonguarders; Brood hiders; Speleophils (rock cavity).
A_2_4C	Binary	Brood hiders; Speleophils (cavity generalist rock crevices, and also under log bark, openings in vegetation, metal cans, etc).
B_1_3A	Binary	Guarders; Substratum choosers; Lithophils.
B_1_4	Binary	Guarders; Substratum choosers; Phytophils.
B_2_2	Binary	Guarders; Nest spawners; Polyphils.
B_2_3A	Binary	Guarders; Nest spawners; Lithophils (rock-gravel).
B_2_3B	Binary	Guarders; Nest spawners; Lithophils (gravel-sand).
B_2_4	Binary	Guarders; Nest spawners; Ariadnophils.
B_2_5	Binary	Guarders; Nest spawners; Phytophils.
B_2_6	Binary	Guarders; Nest spawners; Psammophils.
B_2_7A	Binary	Guarders; Nest spawners; Speleophils (rock cavity/roof).
B_2_7B	Binary	Guarders; Nest spawners; Speleophils (bottom burrows or natural holes associated with structure or bank).
B_2_7C	Binary	Guarders; Nest spawners; Speleophils (cavity generalist).
C1_3_4_C2_4	Binary	A lumping of all bearers. May also be regarded as substrate-indifferent.
MAXTL	Continuous	Maximum total length in centimeters. Some records may estimate asymptotic length. We assumed that the difference between maximum observed length and asymptotic length is negligible.
MATUAGE	Continuous	Mean, median, or modal age at maturity in years for females. Where different ages at maturity were obtained for distinct populations, the clearly supported records were averaged. Male maturity age was accepted where female data were not available.

LONGEVITY	Continuous	Longevity in years based on life in the wild wherever available. Where not indicated, the record was assumed to be from the wild. If wild records were not known, a record from captivity was accepted.
FECUNDITY	Count	Maximum reported fecundity.
SERIAL	Binary	Serial or batch spawner, as inferred from evidence in any part of the species range. Accepted evidence included observations that multiple clutches of eggs are released over the spawning season by an individual female or that distinct batches of eggs at different stages of maturation are in the ovaries.
JAN–DEC	Continuous	The proportion of the month in the species' spawning season, following the convention of Menhinick (1991). Data were converted from weeks to proportion of the month, yielding precision of 0.25. See also SEASON and REPSTATE.
SEASON	Continuous	The sum of the proportions of each month in which spawning occurs; approximates length of the spawning season.
REPSTATE	Categorical	Representative state (or region) whose account was used to record spawning season. When the account specified the source of the spawning season information, we placed the record as close as possible to its true geographic origin. Like many traits, spawning season length is correlated with temperature, so some reference to geographic location is useful. The codes are the two-letter state abbreviations. Unconventional codes included CAN (Canada), MEX (Mexico), NES (northeastern states), Great Lakes, L. Huron (Lake Huron), and L. Superior (Lake Superior).

Habitat Preference Traits

MUCK	Binary	Muck substrate.
CLAYSILT	Binary	Clay or silt substrate.
SAND	Binary	Sand substrate.
GRAVEL	Binary	Gravel substrate.
COBBLE	Binary	Cobble or pebble substrate.
BOULDER	Binary	Boulder substrate.

BEDROCK	Binary	Bedrock substrate.
VEGETAT	Binary	Aquatic vegetation.
DEBRDETR	Binary	Organic debris or detrital substrate.
LWD	Binary	Large woody debris.
PELAGIC	Binary	Open water.
PREFLOT	Binary	Lotic and lentic systems but more often in lotic.
PREFLEN	Binary	Lotic and lentic systems but more often in lentic.
LARGERIV	Binary	Medium to large river.
SMALLRIV	Binary	Stream to small river.
CREEK	Binary	Creek.
SPRGSUBT	Binary	Spring or subterranean water.
LACUSTRINE	Binary	Lentic systems.
POTANADR	Binary	Potamodromous or anadromous. Species that exhibit significant movement related to spawning. We concentrated on movements between marine and freshwater or within freshwater from large river, reservoirs, or lakes to tributary streams.
LOWLAND	Binary	Lowland elevation.
UPLAND	Binary	Highland elevation.
MONTANE	Binary	Mountainous physiography.
SLOWCURR	Binary	Slow current.
MODCURR	Binary	Moderate current.
FASTCURR	Binary	Fast current.
Salinity and Temperature Tolerances		
EURYHALINE	Binary	Species with wide salinity tolerance; mostly obtained from species zoogeographic summaries in Hocutt and Wiley (1986).
MINTEMP	Continuous	The 30-year average minimum January temperature at range centroid in degrees Celsius. Range centroids were used to extract values from 400-m resolution temperature grids obtained from Climate Source, Inc., Oregon.

MAXTEMP	Continuous	30-year average maximum July temperature at range centroid in degrees Celsius. Range centroids were used to extract values from 400-m resolution temperature grids obtained from Climate Source, Inc., Oregon.
Geographic Distribution		
AREA	Continuous	Range area in square kilometers.
PERIMETER	Continuous	Range perimeter in kilometers.
PATCHES	Count	Number of separate patches in the distribution map. This metric is sensitive to map resolution. We assumed that the relative patchiness of species ranges is preserved when maps are compared at the same resolution. This assumption has not been rigorously verified.
LATRANGE	Continuous	Latitudinal range of species in kilometers. This metric may have utility in conjunction with temperatures at range centroids in estimating species temperature tolerances.
LONRANGE	Continuous	Longitudinal range of species in kilometers.
LONGCENTROID	Continuous	Longitude at range centroid in decimal degrees.
LATCENTROID	Continuous	Latitude at range centroid in decimal degrees.
Conservation Status		
LISTED	Binary	Listed as endangered, threatened, vulnerable, or a subspecies is listed as any one of these categories.
REASON	Categorical	Endangered (1), Threatened (2), Vulnerable (3), or subspecies listed as any of these categories (4).
LIST1	Binary	Present or threatened destruction, modification, or reduction of habitat or range.
LIST2	Binary	Over-exploitation for commercial, recreational, scientific, or educational purposes, including international eradication via indirect impacts of fishing.
LIST3	Binary	Disease or parasitism.
LIST4	Binary	Other natural or anthropogenic factors that affect a taxon's existence, including impacts of nonindeginous organisms, hybridization, competition and/or predation.
LIST5	Binary	Restricted range.

EXTINCT	Categorical	Classified as extinct under one of the extinction categories in Harrison and Stiassny (1999). The following categories are represented: Resolved Extinction ─1500-1948 (A), Unresolved Extinction ─Extinctions since 1948 but No Other Ambiguity (C), Taxonomy Substantiated but Effective Extinction Date Unsubstantiated by Sampling (D2), Poorly Defined Extinctions with Taxonomy Substantiated but Effective Extinction Date Unsubstantiated by Population Decline or Environmental Threat (D3), and Species extinction status unclassified but there is reason to further investigate possible extinction (U).