

# **DOCUMENTING TAXONOMIC DATA QUALITY FOR FISH FIELD IDENTIFICATIONS**

## **A PROPOSAL FOR NATIONAL SURVEYS**

---

*J.B. Stribling<sup>1</sup>, J.E. Flotemersch<sup>2</sup>, R.M. Hughes<sup>3</sup>, L. Reynolds<sup>4</sup>, T. Smith<sup>5</sup>, B.D. Snyder<sup>1</sup>, E. Tarquinio<sup>5</sup>, and C.O. Yoder<sup>6</sup>*

<sup>1</sup>Tetra Tech, Inc., Owings Mills, MD; <sup>2</sup>USEPA/ORD, Cincinnati, OH; <sup>3</sup>Oregon State University, Corvallis, OR; <sup>4</sup>USEPA/Region 3 Aquatic Biology Laboratory, Wheeling, WV; <sup>5</sup>USEPA/OWOW, Washington, DC; <sup>6</sup>Midwest Biodiversity Institute, Columbus, OH

# Acknowledgement

---

## **Fish indicators workgroup, National Rivers and Streams Assessment (NRSA)**

**Lou Reynolds, Chair**

**USEPA Region 3/Aquatic Biology Lab, Wheeling, WV**

**Workgroup members: Leroy Kleinsasser, Brian Hill, Brooks Trammel, Chris Millard, Cody Jones, Dave Bolgrien, Dennis McCauley, Erich Emery, Eric Eiseminger, Chris Yoder, Joe Flotemersch, Lil Herger, Treda Smith, Ellen Tarquinio, Shannon Hubler, Jason Childress, John Wirts, Leanne Stahl, Mark Pearson, Mike Feist, Bob Hughes, Mike Meador, Ron Klauda, Rick Spear, Sam Stribling, Hillary Snook, Susan Holdsworth, Bill Botts**

# **What does “taxonomic data quality” mean?**

---

**Data quality is the magnitude of error associated with a particular dataset  
(Taylor 1988)**

**Taxonomic data quality, then, is the magnitude of error associated with a taxonomic dataset**

# **National Rivers and Streams Assessment (NRSA)**

---

- **1,800 sites over 2-year period (2008-09)**
- **200 repeat samples**
- **Fish, fish tissue, benthic macroinvertebrates, periphyton, water chemistry, physical habitat**
- **Probability-based site selection**

# National Rivers and Streams Assessment (NRSA)

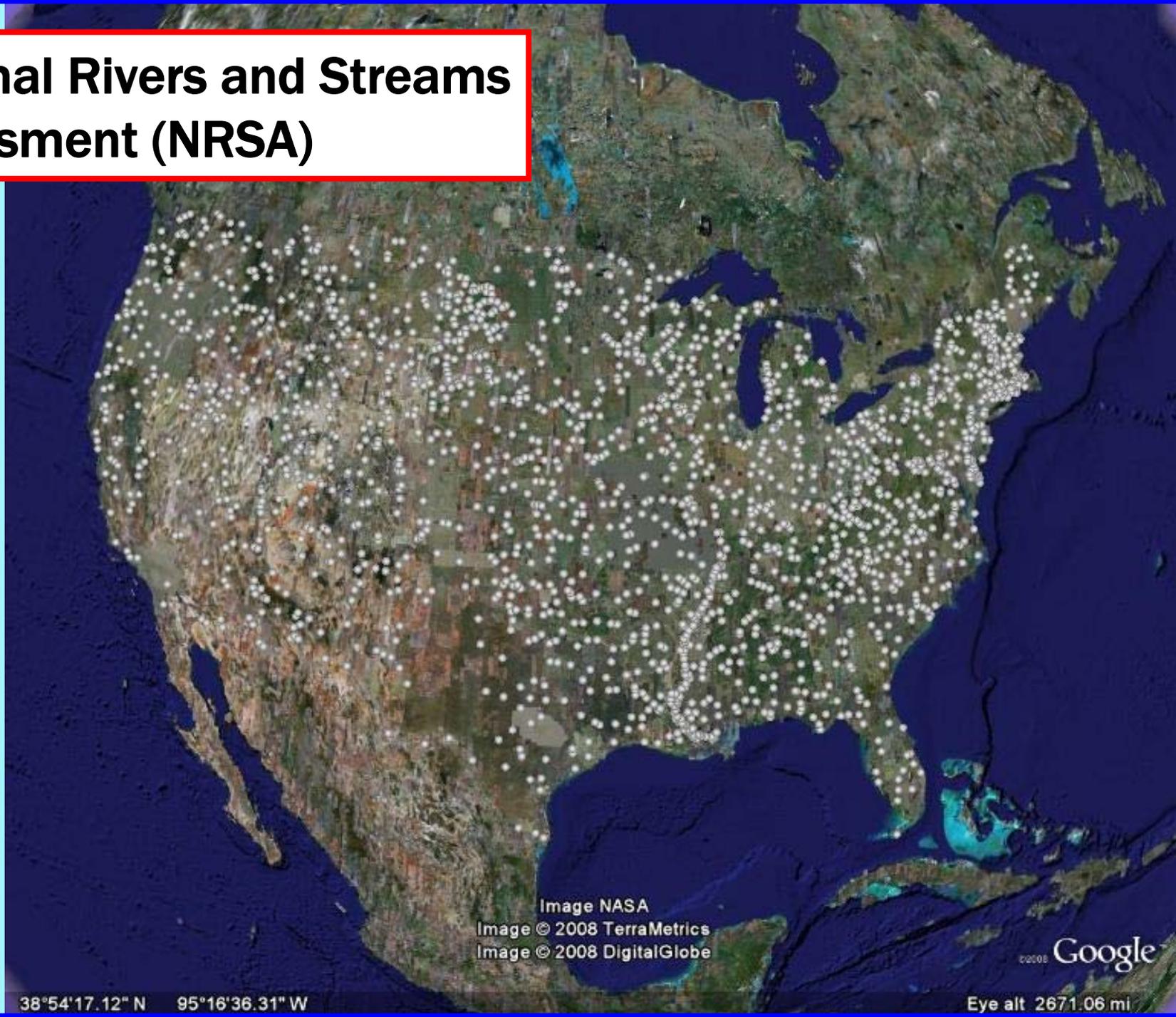


Image NASA  
Image © 2008 TerraMetrics  
Image © 2008 DigitalGlobe

©2008 Google

38°54'17.12" N 95°16'36.31" W

Eye alt 2671.06 mi

# Taxonomic QC Goal

---

- **What we can say with the results (examples):**
  - **“The taxonomic precision associated with the NRSA fish dataset is xx%”**
  - **“Field team X with Taxonomist Y had an identification error rate of xx”**
  - **“The striped jumprock seemed to be consistently called the brassy jumprock in region Z, and Field team X with Taxonomist Y did all of that region. Need to make corrections to dataset”**







<b>Species</b>	<b>Count</b>
Altamaha shiner	2
Black crappie	1
Blackbanded darter	26
Bluefin stoneroller	25
Bluegill	51
Bluehead chub	33
Brassy jumprock	1
Creek chub	1
Flat bullhead	114
Green sunfish	3
Largemouth bass	1
Redbreast sunfish	124
Rosyface chub	8
Snail bullhead	7
Spottail shiner	19
Striped jumprock	3
Turquoise darter	18
Yellow bullhead	1
Yellowfin shiner	7

**Ocmulgee River watershed  
(Georgia, USA), Snapfinger  
Creek (Station ID\_SN8),  
drainage area 43mi<sup>2</sup>**

**Dr. Chris Skelton**

# Sample voucher

---

- **Voucher frequency and distribution**
  - **1 complete voucher for every 10 sites**
  - **At least 1 voucher per taxonomist**
- **Preserved specimens, and**
- **Photovouchers (digital images)**
  - **Rare, threatened, and endangered species**
  - **Large-bodied species/individuals**
  - **(Very common species)**

# Preserved specimens

---

- **Enough specimens per species (at least 5, no more than 20) to allow recognition**
- **Place in single container**
- **Label (site number, drainage basin, date)**
- **Intended to be 'blind samples'**

# QC process

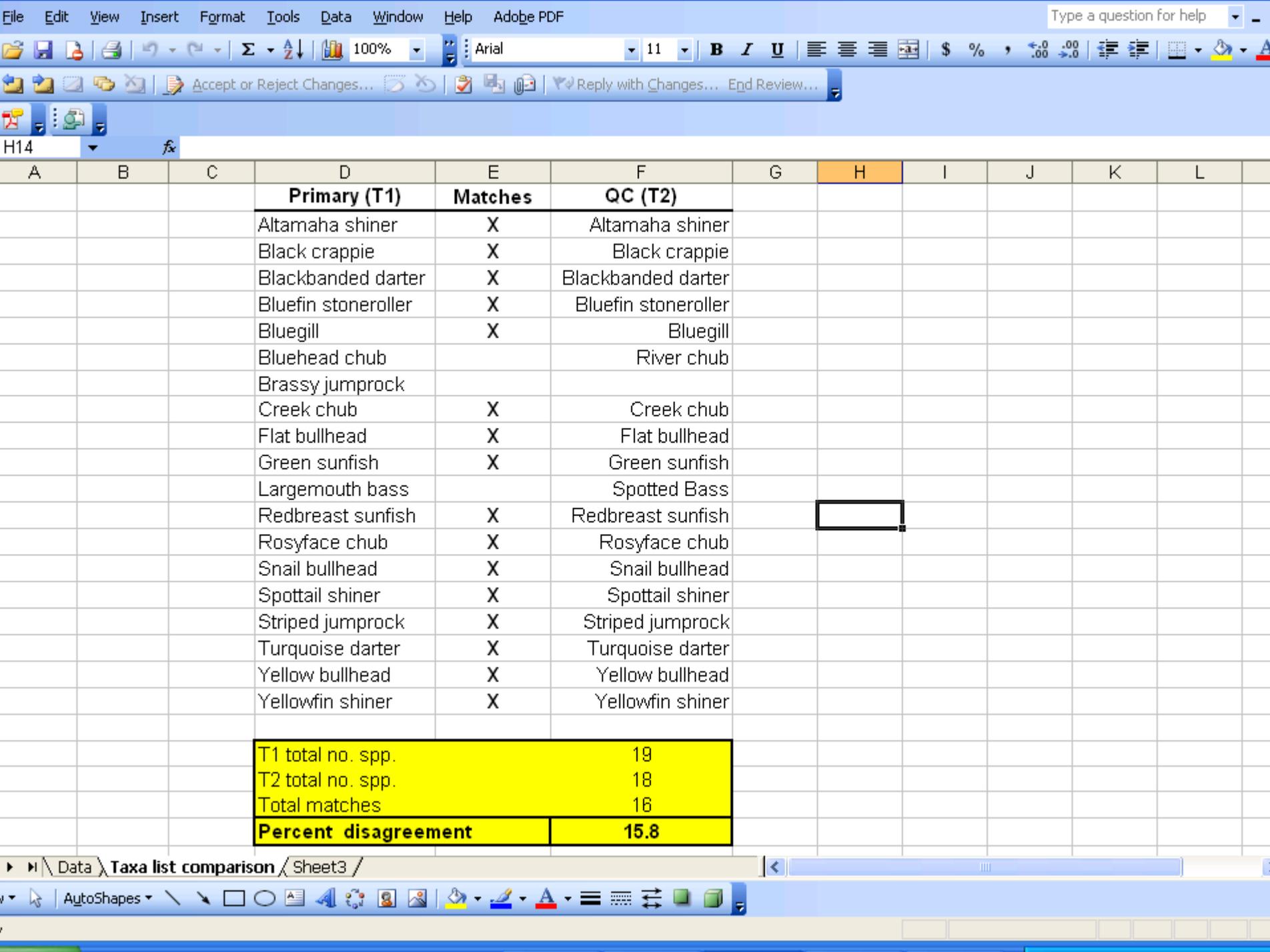
---

- **Regional experts will reconstruct species list for all sample vouchers**
- **Match species lists from 2 taxonomists for same sample**
  - **Primary data – field identifications supplemented by lab confirmation**
  - **QC/sample voucher data**
- **Quantify differences in 2 lists**

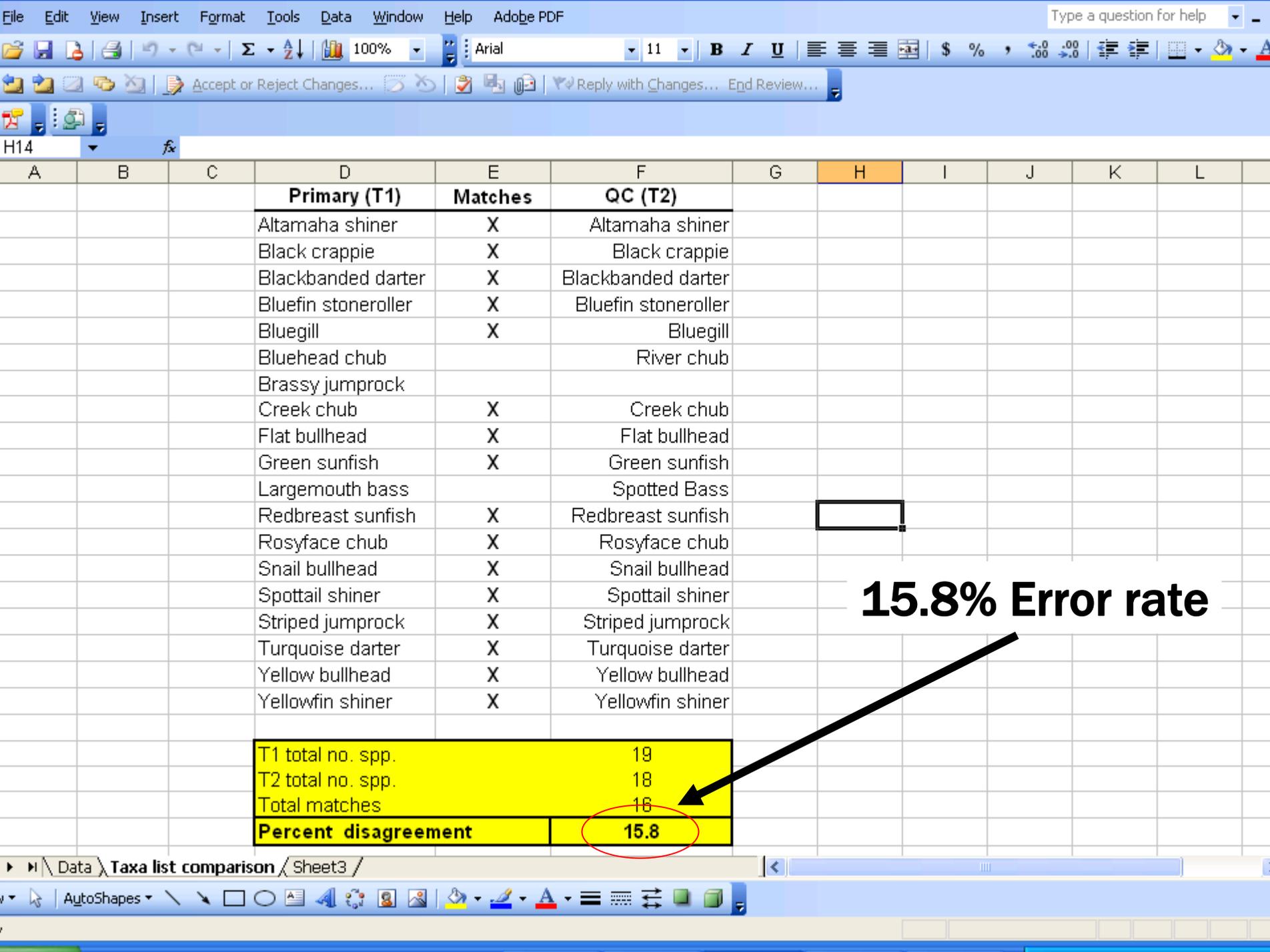
# **Voucher repositories**

---

- **University of Georgia, Institute of Ecology**
- **University of Texas, Museum of Natural History**
- **Ohio State University/Museum of Biodiversity/Midwest Biodiversity Institute**
- **(University of Oklahoma)**



	A	B	C	D	E	F	G	H	I	J	K	L
				<b>Primary (T1)</b>	<b>Matches</b>	<b>QC (T2)</b>						
				Altamaha shiner	X	Altamaha shiner						
				Black crappie	X	Black crappie						
				Blackbanded darter	X	Blackbanded darter						
				Bluefin stoneroller	X	Bluefin stoneroller						
				Bluegill	X	Bluegill						
				Bluehead chub		River chub						
				Brassy jumprock								
				Creek chub	X	Creek chub						
				Flat bullhead	X	Flat bullhead						
				Green sunfish	X	Green sunfish						
				Largemouth bass		Spotted Bass						
				Redbreast sunfish	X	Redbreast sunfish						
				Rosyface chub	X	Rosyface chub						
				Snail bullhead	X	Snail bullhead						
				Spottail shiner	X	Spottail shiner						
				Striped jumprock	X	Striped jumprock						
				Turquoise darter	X	Turquoise darter						
				Yellow bullhead	X	Yellow bullhead						
				Yellowfin shiner	X	Yellowfin shiner						
				<b>T1 total no. spp.</b>		19						
				<b>T2 total no. spp.</b>		18						
				<b>Total matches</b>		16						
				<b>Percent disagreement</b>		15.8						



	A	B	C	D	E	F	G	H	I	J	K	L
				Primary (T1)	Matches	QC (T2)						
				Altamaha shiner	X	Altamaha shiner						
				Black crappie	X	Black crappie						
				Blackbanded darter	X	Blackbanded darter						
				Bluefin stoneroller	X	Bluefin stoneroller						
				Bluegill	X	Bluegill						
				Bluehead chub		River chub						
				Brassy jumprock								
				Creek chub	X	Creek chub						
				Flat bullhead	X	Flat bullhead						
				Green sunfish	X	Green sunfish						
				Largemouth bass		Spotted Bass						
				Redbreast sunfish	X	Redbreast sunfish						
				Rosyface chub	X	Rosyface chub						
				Snail bullhead	X	Snail bullhead						
				Spottail shiner	X	Spottail shiner						
				Striped jumprock	X	Striped jumprock						
				Turquoise darter	X	Turquoise darter						
				Yellow bullhead	X	Yellow bullhead						
				Yellowfin shiner	X	Yellowfin shiner						
				<b>T1 total no. spp.</b>		19						
				<b>T2 total no. spp.</b>		18						
				<b>Total matches</b>		16						
				<b>Percent disagreement</b>		15.8						

**15.8% Error rate**



# Potential pitfalls

---

- **Field teams don't voucher properly and objectively**
- **Condition of sample vouchers making re-identifications problematic**
  - **Specimens degraded**
  - **Photovouchers inadequate for key characters**
- **Inappropriate regional experts**

# **...Stay tuned**

---

- **Questions?**