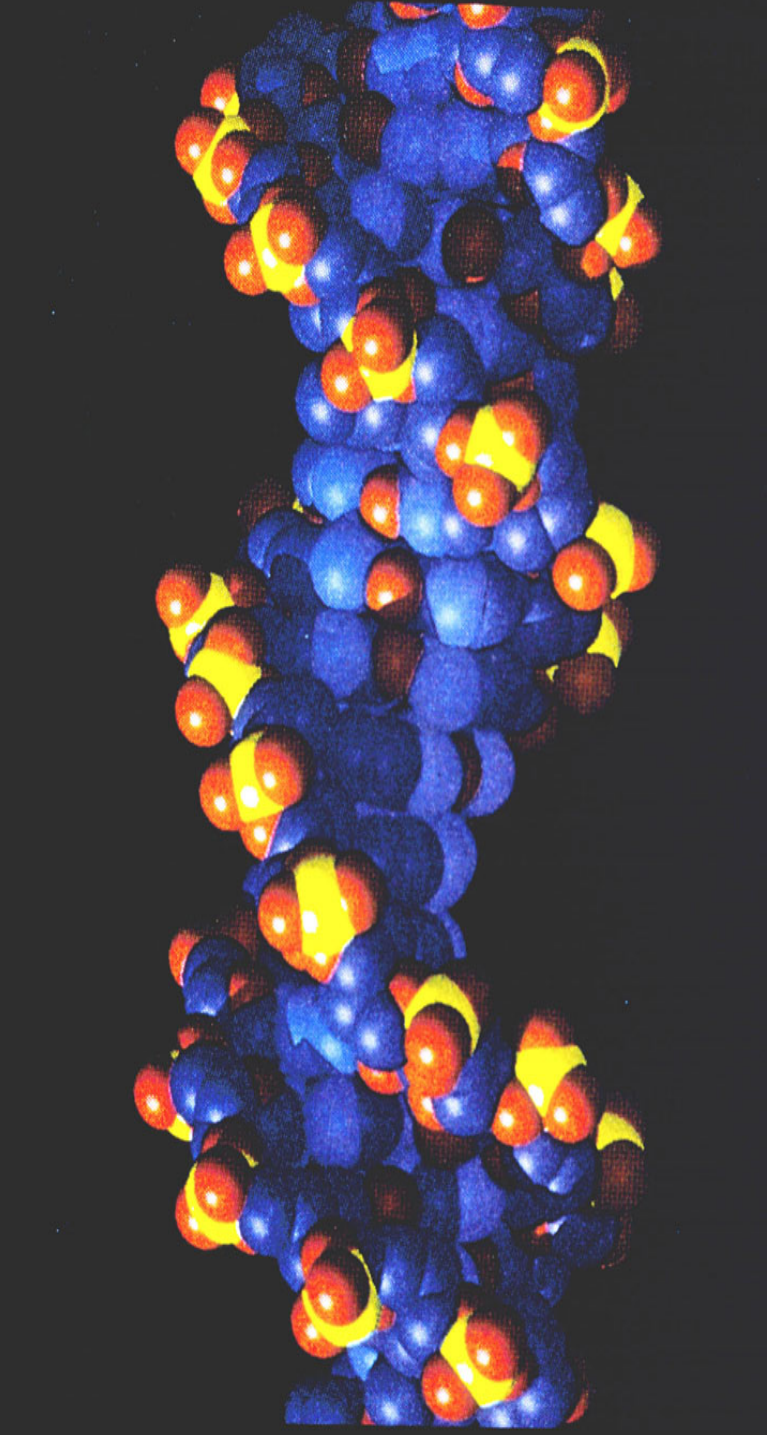


Databases for Genetic Stock Identification:

Overview and Historical Perspective

Lisa Seeb and Bill Templin

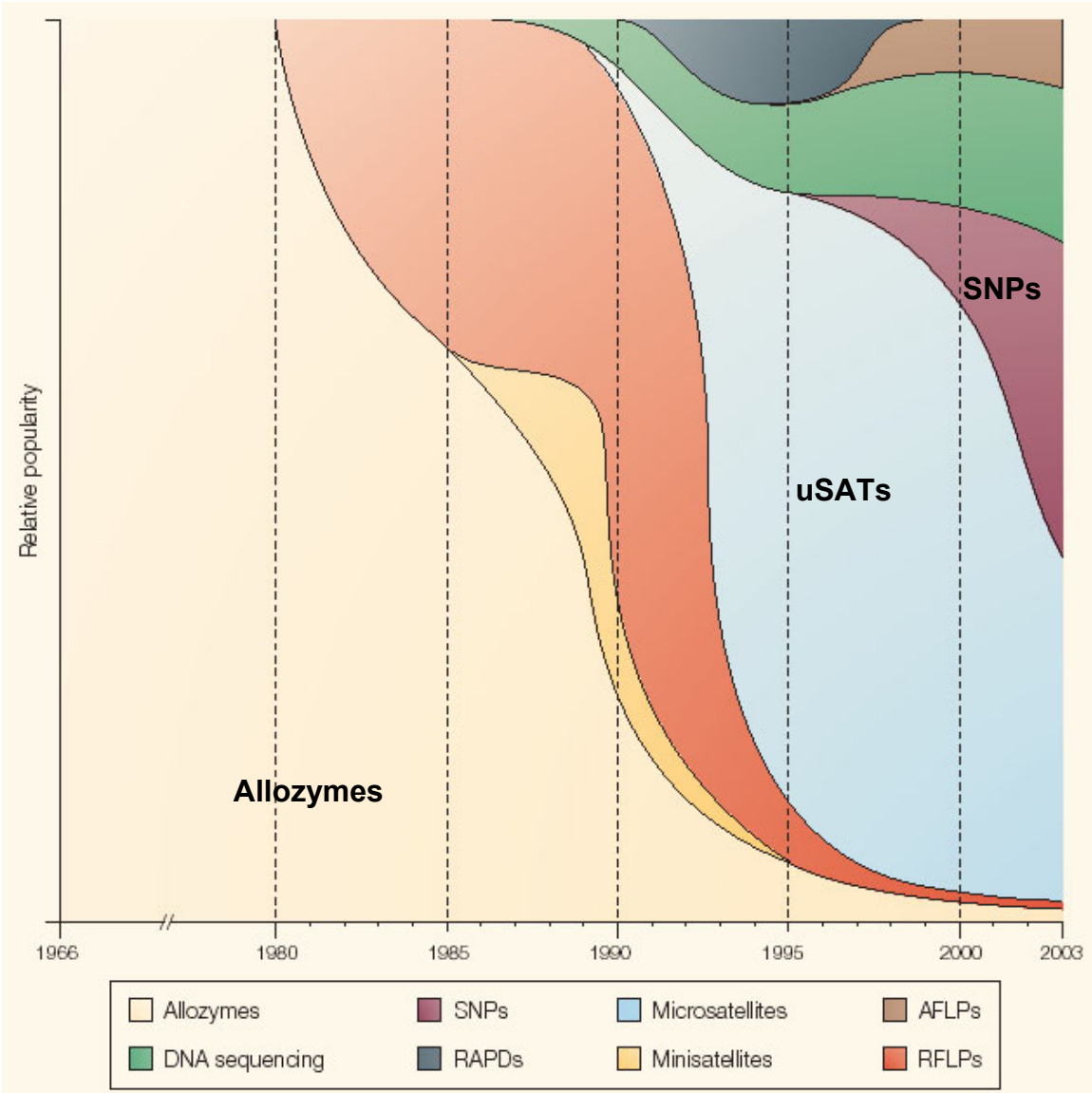
Gene Conservation Laboratory

If genetic markers are to be used for replacement or augmentation of the existing coastwide CWT program, then important considerations are:

- Resolution
  - Regional, fine scale, or both
- Coverage
  - Geographic
  - Single laboratory or collaborative effort
- Cost/speed
- Standardization among labs
  - economy for migratory and mixture studies
  - transparency of data for multi-jurisdictional management
- Expandability
  - Add populations
  - Add loci, archive reference material

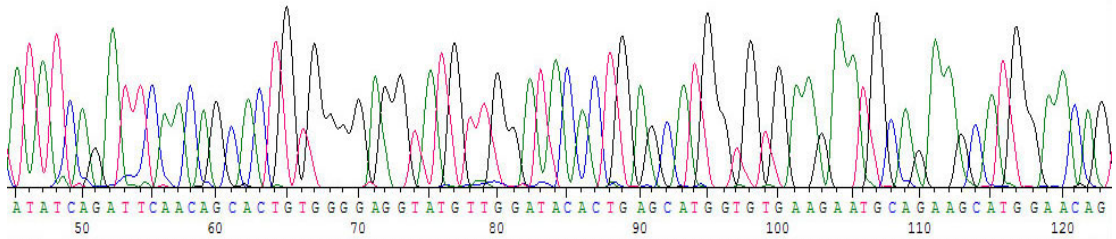
Relative Popularity of Markers

- Resolution
- Coverage
- Standardization
- Cost/Speed
- Expandability



Schlötterer 2004

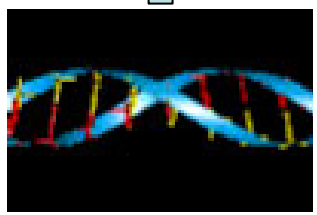
DNA Sequence



- Highest resolution
- Least ambiguity (reproducible in any lab)

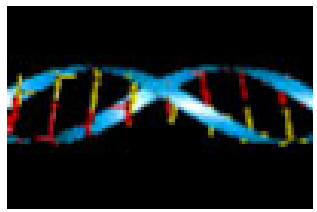
But…

- Relatively time-consuming / costly



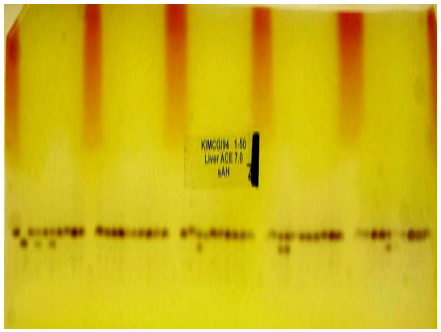
DNA

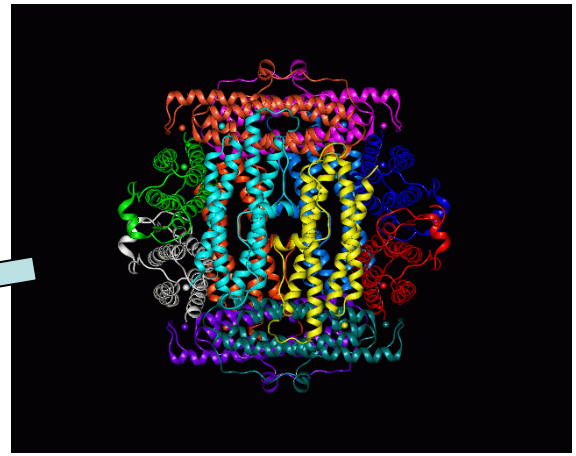
Genetic markers are tools for learning something about DNA sequence without spending the time or money to do the sequencing. The savings in time and money come at a cost of resolution or certainty in data or both.



DNA

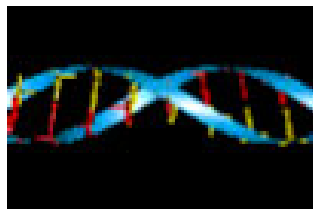
allozymes





protein

RNA



DNA

First Chinook Salmon Baselines

- Late 1970's: Target Fisheries: Columbia River
  24 Columbia Basin stocks
  10 loci
  One lab (NMFS)

- Early 1980's: Target Fisheries: WA - OR Coastal
  49 CA to BC stocks
  14 loci
  One lab (NMFS)

- 1983: Target Fisheries: CA - OR - WA Coastal
  Two Labs (NMFS and UCD)
  Two Baselines
  (one group of fishery managers)

1984 –present
Multiple Labs - One Chinook Salmon Baseline

Common set of standardized loci and alleles

Methods of baseline construction

Periodic updates

Fishery analysis methods

Data use agreements

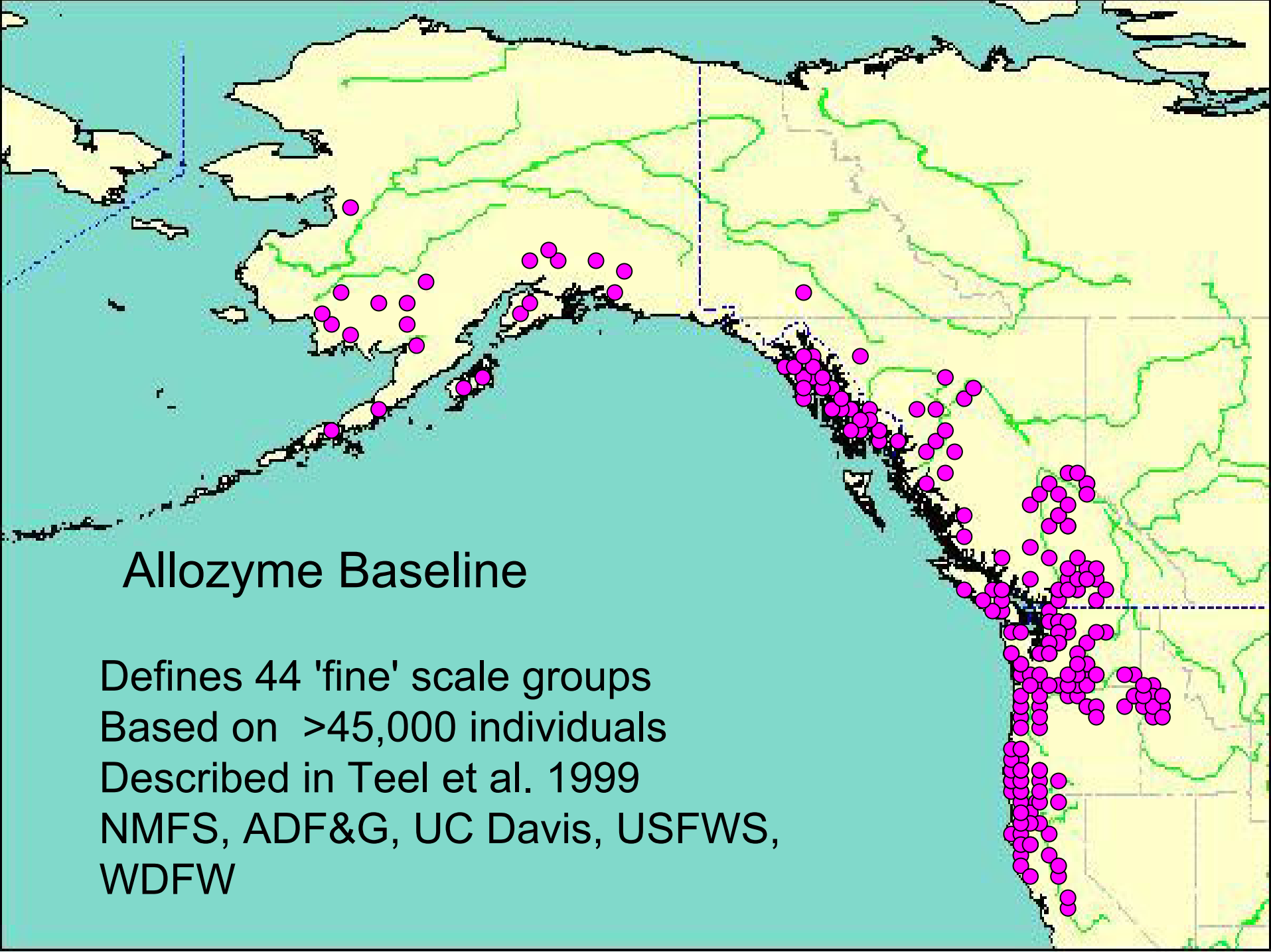
Documentation

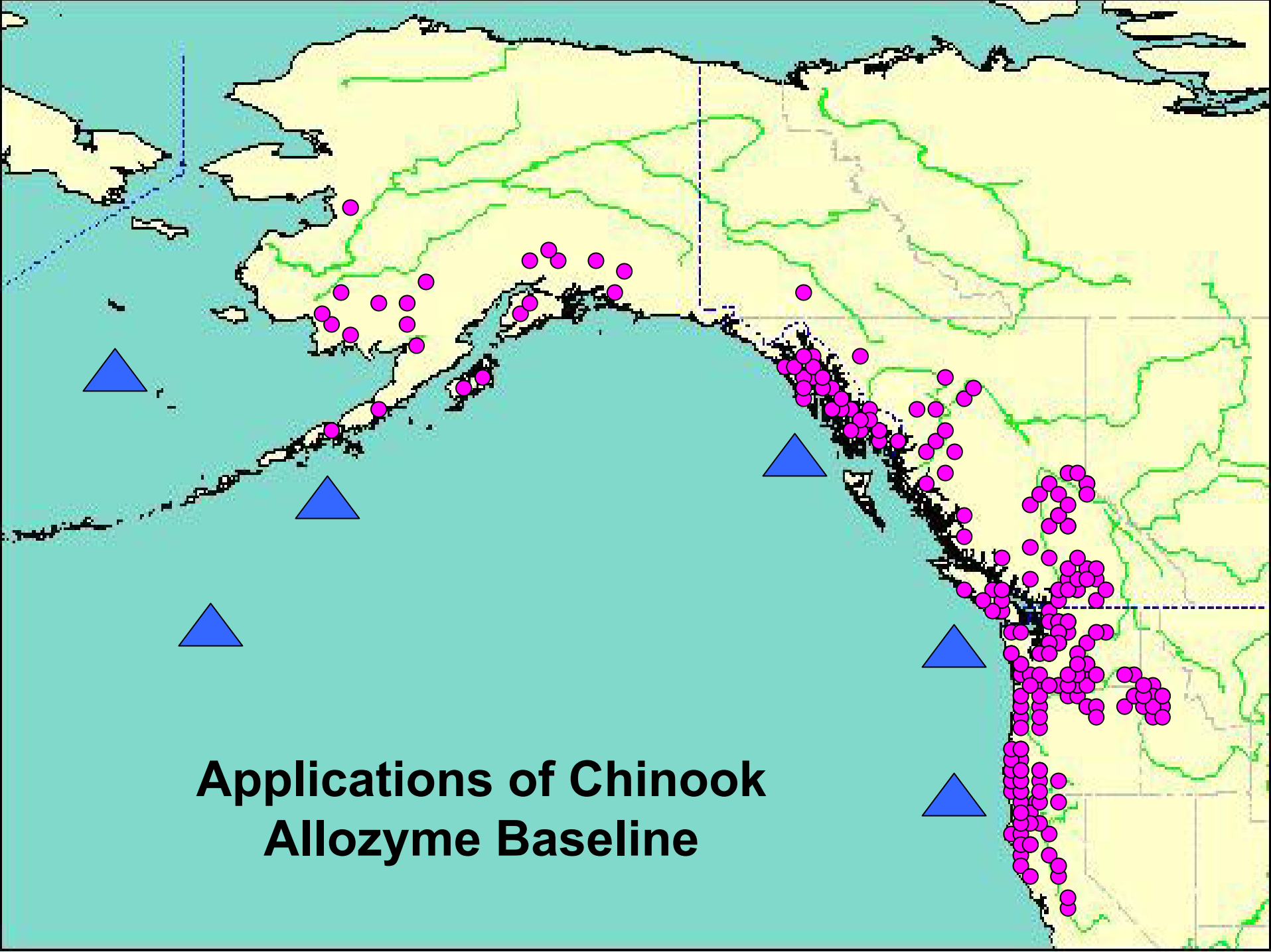
see Shaklee and Phelps 1990

NMFS / UCD Collaborations 1984
What to Share?

- Samples? 2nd hand tissues
  tissue duplicates
  = two baselines

- Data? different sets of loci
  different genetic models
  different alleles recognized
  different nomenclature
  = corrupt baseline


Allozyme Baseline
Defines 44 'fine' scale groups
Based on >45,000 individuals
Described in Teel et al. 1999
NMFS, ADF&G, UC Davis, USFWS,
WDFW


Applications of Chinook
Allozyme Baseline

Database Attributes - Chinook Salmon

	Allozymes	uSATS	SNPs
Resolution	Regional +		
Coverage (collaborative or single)	Pacific Rim, collaborative		
Standardization	Required		
-Time/cost	Complete		
-Markers	33 loci		
-Populations	255		
Cost/fish	$40 ↑		
Expandability			
-Populations	Unlikely		
-Loci	None		

Allozyme Database Limitations

- Expandability very limited
  - No archival reference tissues
  - Allozyme loci limited
- Geographically-specific locus sets
  - Critical loci missing from Russia, poor sample quality
  - Yukon River samples discarded

DNA Developments

1990's – present

Chinook Salmon

- Advantages
  - Many techniques,"unlimited" loci
    - Sequencing, RFLPs, minisatellites, microsatellites, AFLPs, SNPs
  - High resolution potential
  - Non lethal sampling
- Disadvantages
  - Many techniques, "unlimited" loci

DNA Developments

1990's – present

"We believe that microsatellites are the most promising candidates among novel marker systems to effectively supplement and perhaps eventually supplant allozyme markers."

Wright and Bentzen (1994)

Microsatellite (uSAT) Attributes

- Advantages
  - High resolution in many areas, individual ID possible
    - Some loci hypervariable >70 alleles
  - Selectively neutral loci
  - Technology in most labs
- Limitations
  - Standardization is time consuming and costly
  - Some loci hypervariable >70 alleles
    - Large sample sizes required, N > 150
    - Statistical considerations

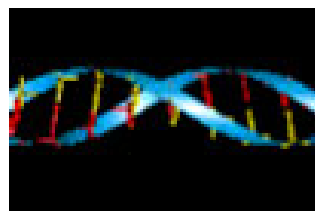
DNA secondary structure

(microsatellite example)

?

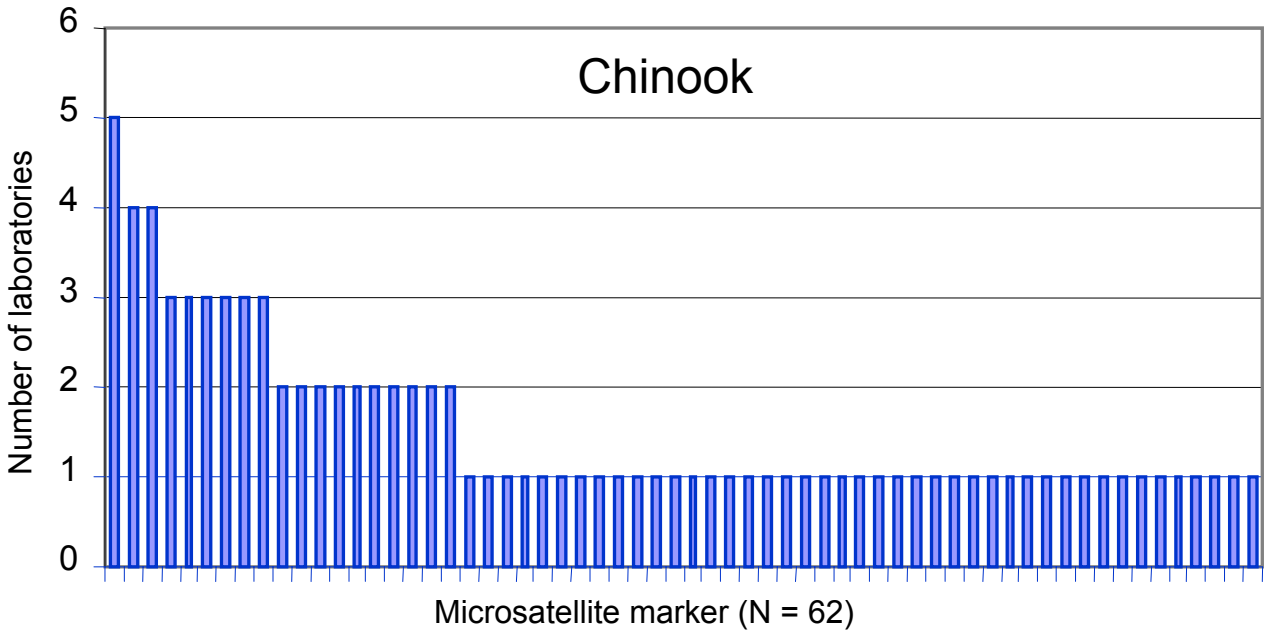
Allele	
1	2
89	93
85	90

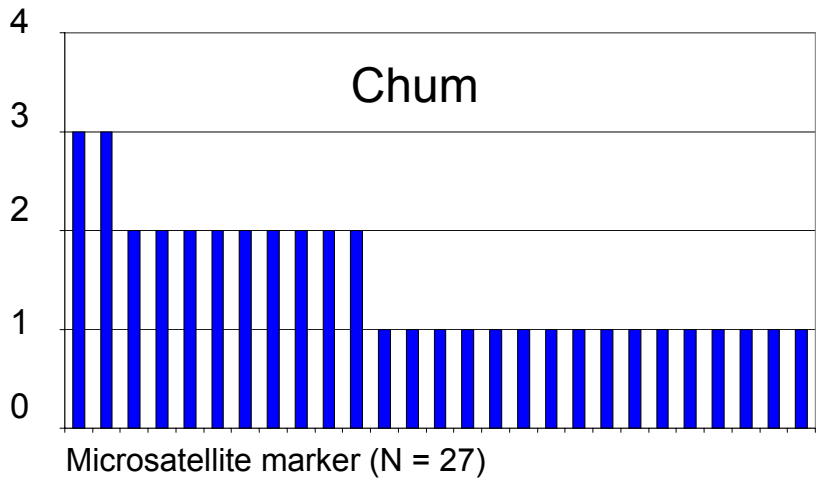
CACACACACACACACACACA …

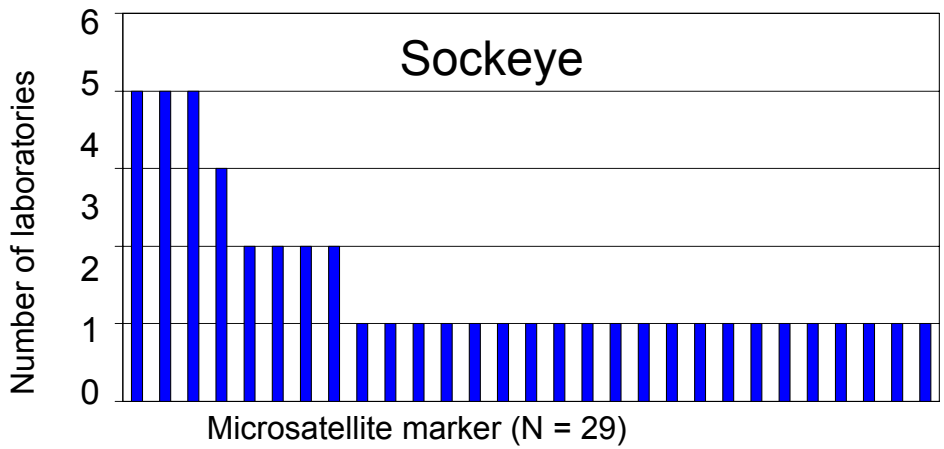


DNA

Microsatellite Standardization







Coastwide uSAT Standardization

- June 1999: Scoping Workshop, Univ. of Montana
  Six species groups organized
  Leads: Chinook (NMFS/NWFSC)
  Coho (CDFO)
  Tasked to develop standardization protocols

- Nov. 2000: Workshop organized by PSC, Vancouver
  Tasked to develop allelic ladders and list of loci
  Discussion of standardization vs. exchange of samples

PSC, Nov. 2000

~~NMFS / UCD~~ Collaborations
What to Share?

- Samples? 2nd hand tissues
  tissue duplicates
  = multiple baselines

- Data? different sets of loci
  different genetic models
  different alleles recognized
  different nomenclature
  = corrupt baseline

Or collaborate?

Coastwide uSAT Standardization

- June 1999: Scoping Workshop, Univ. of Montana
  Six species groups organized
  Leads: Chinook (NMFS/NWFSC)
  Coho (CDFO)
  Tasked to develop standardization protocols

- Nov. 2000: Workshop organized by PSC, Vancouver
  Tasked to develop allelic ladders and list of loci
  Discussion of standardization vs. exchange of samples

- August 2001: Workshop organized by UC Davis
  Chinook group presented results for successful standardization of single locus, *Ots2*

Consideration of markers not requiring standardization began

Coastwide uSAT Standardization

- June 1999: Scoping Workshop, Univ. of Montana
  Six species groups organized
  Leads: Chinook (NMFS/NWFSC)
  Coho (CDFO)
  Tasked to develop standardization protocols

- Nov. 2000: Workshop organized by PSC, Vancouver
  Tasked to develop allelic ladders and list of loci
  Discussion of standardization vs. exchange of samples

- August 2001: Workshop organized by UC Davis
  Chinook group presented results for successful standardization of single locus, *Ots2*

- Nov. 2002: Workshop organized by CTC, Vancouver
  RFP issued for development of coastwide database for Chinook salmon

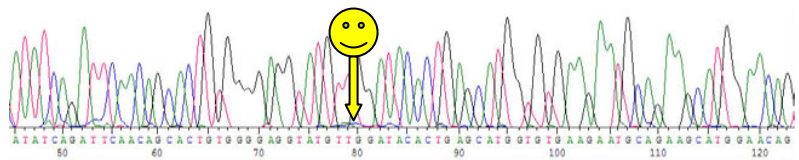
Database Attributes - Chinook Salmon

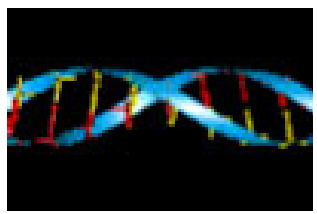
	Allozymes	uSATS	SNPs
Resolution	Regional +	Regional, fine	
Coverage (collaborative or single)	Pacific Rim, collaborative	SE AK – CA, both	
Standardization	Required	Required	
-Time/cost	Complete	1-2 years,>$.5 M 7 labs	
-Markers	33 loci	12-15 loci	
-Populations	255	105	
Cost/fish	$40 ↑	$35 ↑↓	
Expandability			
-Populations	Unlikely	Planned	
-Loci	None	Archived tissue	

DNA sequence detection

(SNP example)

A, C, G, T or -





DNA

SNP Attributes

- Advantages
  - No standardization required, reproducible by definition
  - Two alleles, ease of scoring
  - Smaller sample sizes (N < 100)
  - Assay both nuclear and mitochondrial loci
  - Selective and neutral loci
- Limitations
  - Two alleles, more loci required than uSATS
  - Fewer labs have high throughput technology

Database Attributes - Chinook Salmon

	Allozymes	uSATS	SNPs
Resolution	Regional +	Regional, fine	Regional, fine(?)
Coverage (collaborative or single)	Pacific Rim, collaborative	SE AK – CA, both	Limited Pacific Rim), collaborative
Standardization	Required	Required	By Definition
-Time/cost	Complete	1-2 years,>$.5 M 7 labs	Lab startup only
-Markers	33 loci	12-15 loci	10 developed, >50 candidate
-Populations	255	105	50 (Alaska primarily)
Cost/fish	$40 ↑	$35 ↑↓	$25/fish ↑↓
Expandability			
-Populations	Unlikely	Planned	Planned
-Loci	None	Archived tissue	Archived tissue

Databases for Genetic Stock Identification

Chinook Salmon

- Allozymes
  - 255 pops, 33 loci
  - Phasing out
- Microsatellites
  - Most regions complete, but unstandardized
  - CTC standardized--SE Alaska to California—to be completed June 2005, 12-15 loci, 105 pops
- SNPs
  - Regional utilization
  - Coastwide developing, >50 candidate SNPs available
  - Could be added to CTC DNA database

Statistical Algorithms

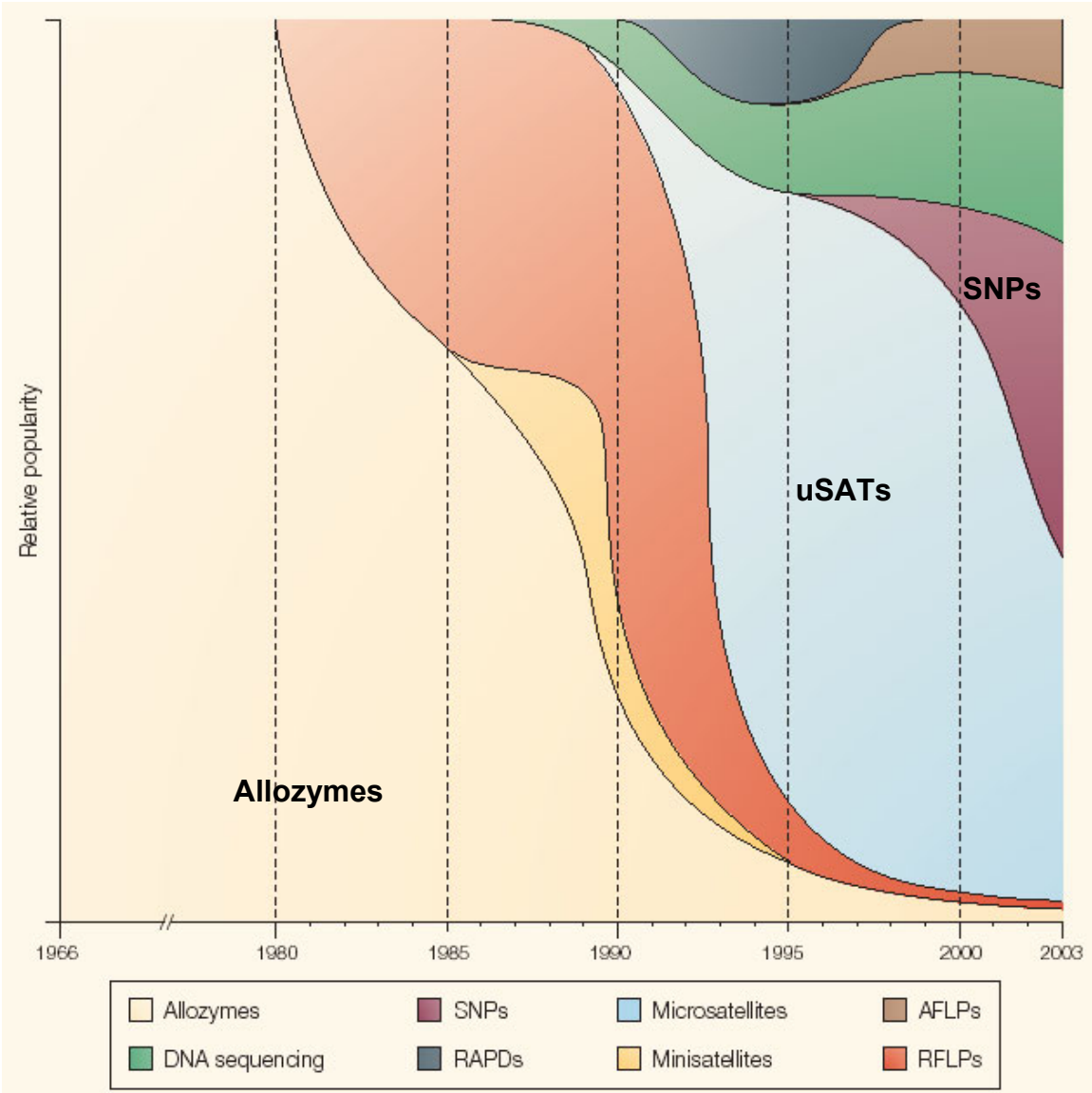
- Estimation
  - Composition
  - Individual assignment
- Algorithms
  - Maximum Likelihood
  - Bayesian

If genetic markers are to be used for replacement or augmentation of the existing coastwide CWT program, then databases must be:

- Standardized
- Interchangeable
- Transparent
- Appropriate coverage for all PSC fisheries
- Expandable
  - Populations
  - Loci

Relative Popularity of Markers

- Resolution
- Coverage
- Standardization
- Cost/Speed
- Expandability



Schlötterer 2004

ADF&G Applications
Chinook Salmon

- Southeast Alaska Fisheries
  - Transitioning from allozymes to CTC Coastwide uSATS and SNPs
- Southcentral Alaska
  - CTC coastwide uSATS and SNPs
- Western Alaska
  - Regional databases now, uSATS and SNPs
  - CTC coastwide in future

SE Alaska Troll Fishery

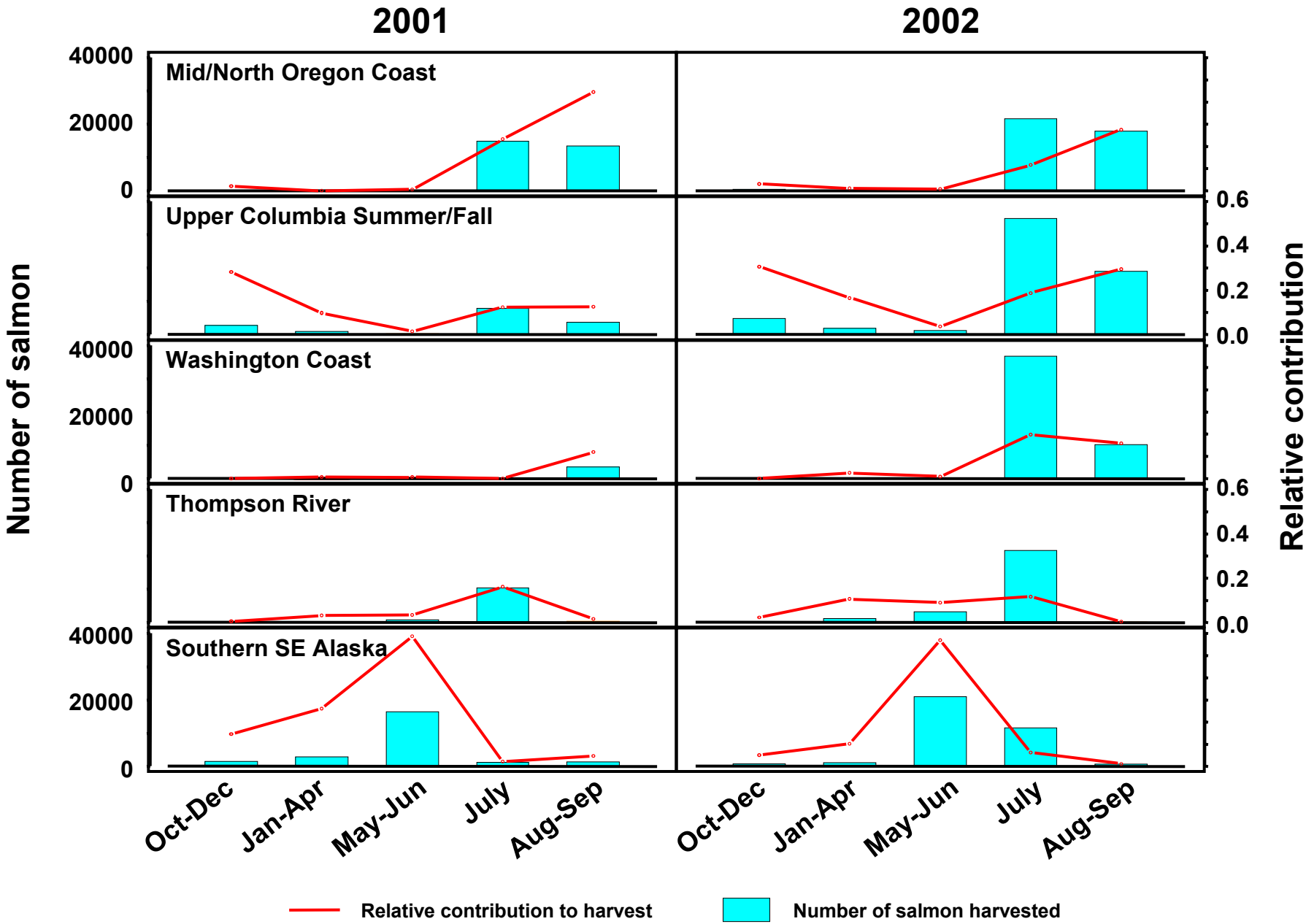
1999-2003

Fishing Seasons:

October-December
January-April 14
April 15-June
July-September

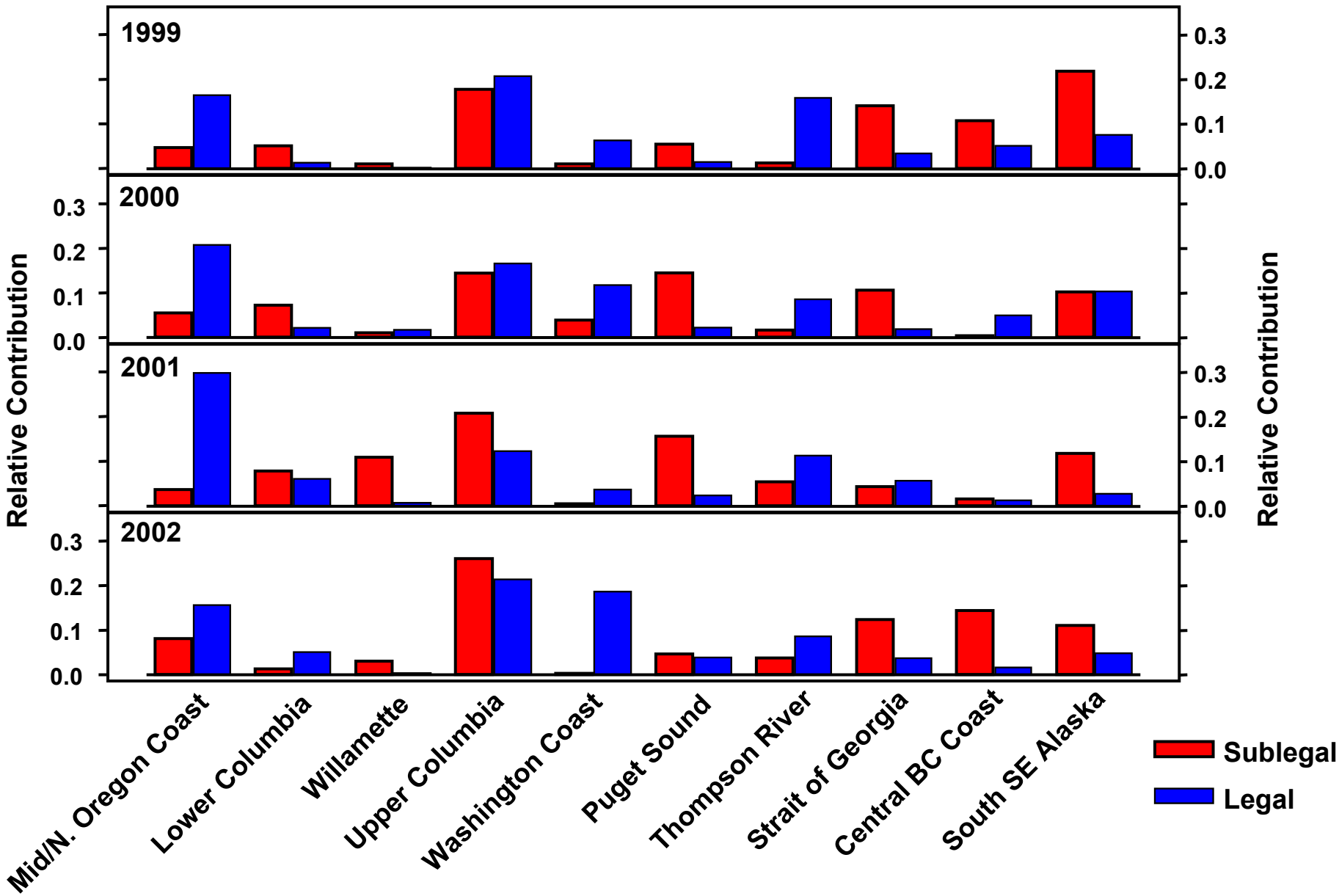
Legal-sized chinook (>28 in)
Sublegal chinook (<28 in)

Annual Variation in Contribution of Selected Stocks



SE Alaska Summer Troll Fishery

Legal and Sublegal

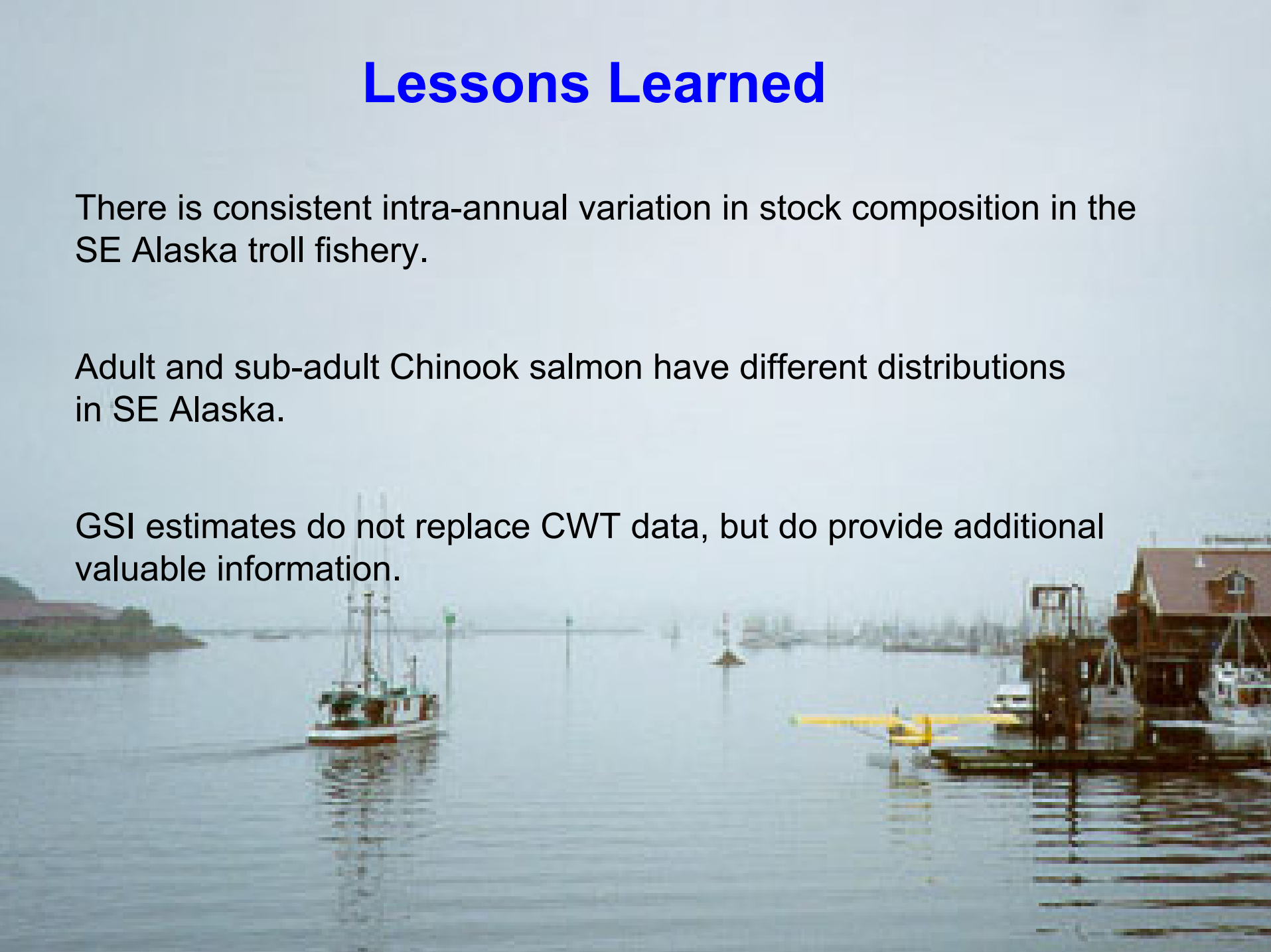


Lessons Learned

There is consistent intra-annual variation in stock composition in the SE Alaska troll fishery.

Adult and sub-adult Chinook salmon have different distributions in SE Alaska.

GSI estimates do not replace CWT data, but do provide additional valuable information.



Beginning 2004

Cover all fisheries involving Chinook salmon

Troll	Gillnet
Seine	Sport

Cover all size-classes (Legal and sublegal)

Include matched scale and size information

Switch to DNA markers for GSI (microsatellites & SNPs)

Annual Schedule for Sampling SE Alaska Chinook Fisheries