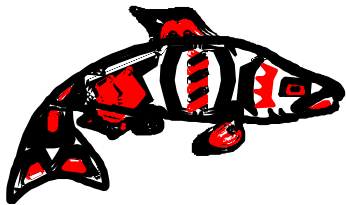# Current affairs in DNA typing for fishery management: From GSI to FPG…marker wars to useful augmentation of CWTs

Pacific Salmon Commission Invited Workshop:
Applications of Genetic Stock Identification (GSI) to Ocean Salmon Management
Portland, Oregon
15 February 2007

Paul Moran

Conservation Biology Division

Northwest Fisheries Science Center

Outline

- A (very) brief history of Genetic Stock Identification and Full Parental Genotyping
- Genetic standardization and the GAPS consortium
- GAPS database and web application
- Opportunities and challenges for DNA methods
- Beyond the controversies: Approaches to facilitate conservation, recovery, and harvest management

History of GSI mixed fishery analysis

- Fred Utter: Salmon allozyme electrophoresis
- George Milner: Columbia R. Chinook GSI, 20 years of high seas and terminal estimates
- Jerry Pella: Further developed analytical methods
- Michael Banks: Microsatellites and Central Valley individual assignment
- Steve Forbes: SNPs, a "new" generation of markers
- Linda Park: Lilliwaup coho pedigree, proof of concept
- Worldwide: Explosion of ecological genetic applications employing microsatellites and increasingly SNPs

How GSI works

- **Baseline**: Known origin, reference population samples likely to contribute to a given mixed fishery sample
- **Mixture**: Fishery sample from unknown contributing populations
- GSI algorithms estimate the proportional contribution from each baseline population
- Populations are combined into "reporting groups"

GSI versus IA

- Genetic Stock Identification—proportional contributions to a mixed fishery

- Individual Assignment of a single fish to its population of origin—a far more demanding proposition
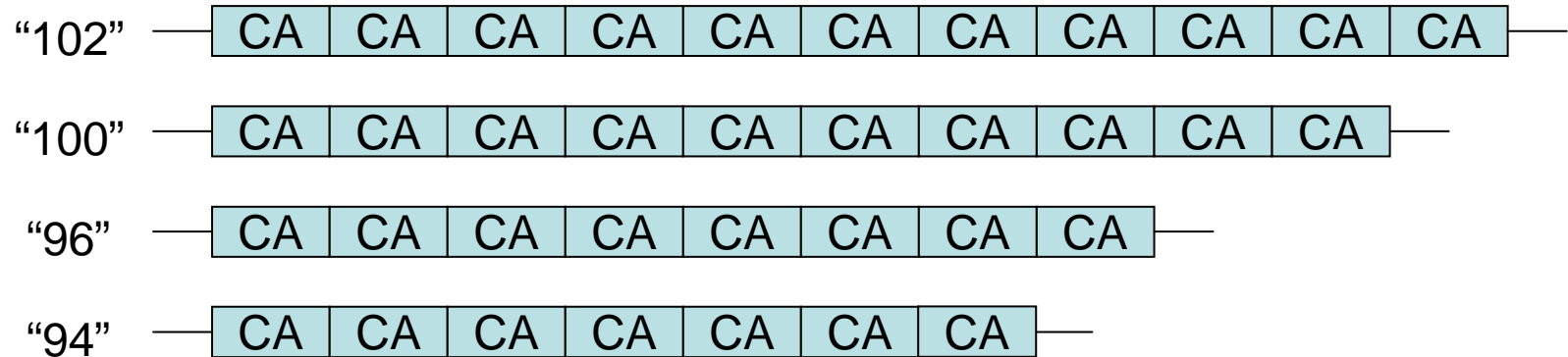
Factors affecting resolution

- Number of baseline populations and number of individuals per population

- Number of individuals in the mixture sample

- Complexity of the mixture

Power of genetic markers (loci)

- Number of independent alleles summed over loci, under ***ideal*** conditions

- Ascertainment bias, preferential targeting of more informative alleles

- Non-neutrality, selection creates larger than average genetic differences (and similarities) among populations

Microsatellite markers—simple sequence repeats
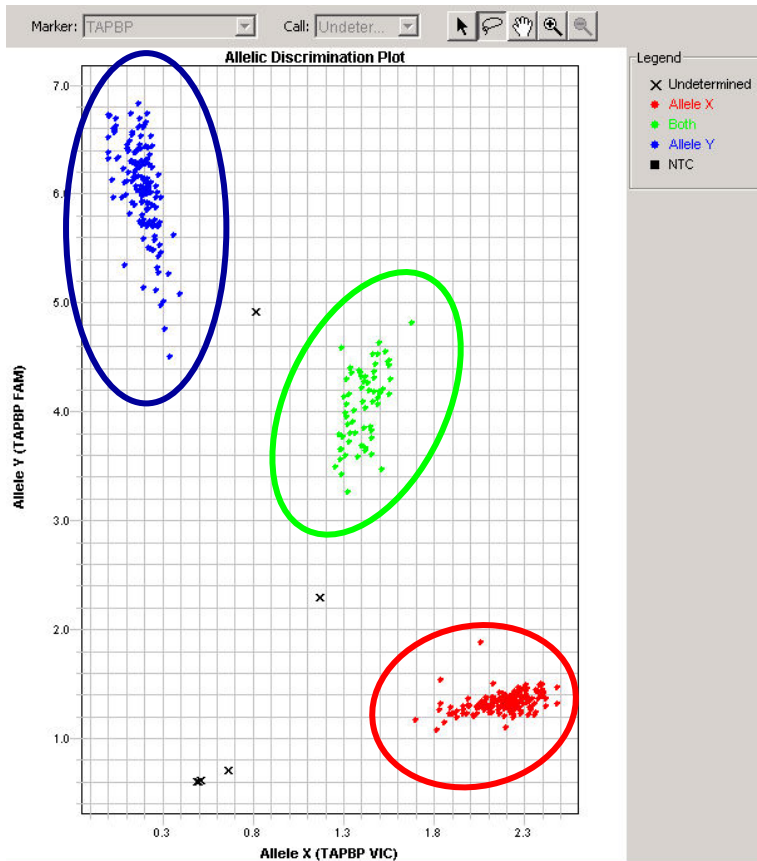
Allele designations typically related to fragment size

"102" — CA | CA | CA | CA | CA | CA | CA | CA | CA | CA | CA

"100" — CA | CA | CA | CA | CA | CA | CA | CA | CA | CA

"96" — CA | CA | CA | CA | CA | CA | CA | CA

"94" — CA | CA | CA | CA | CA | CA | CA

SNPs—point mutations

Simpler allele designations

"T" — GTCGG**T**CACCCTG
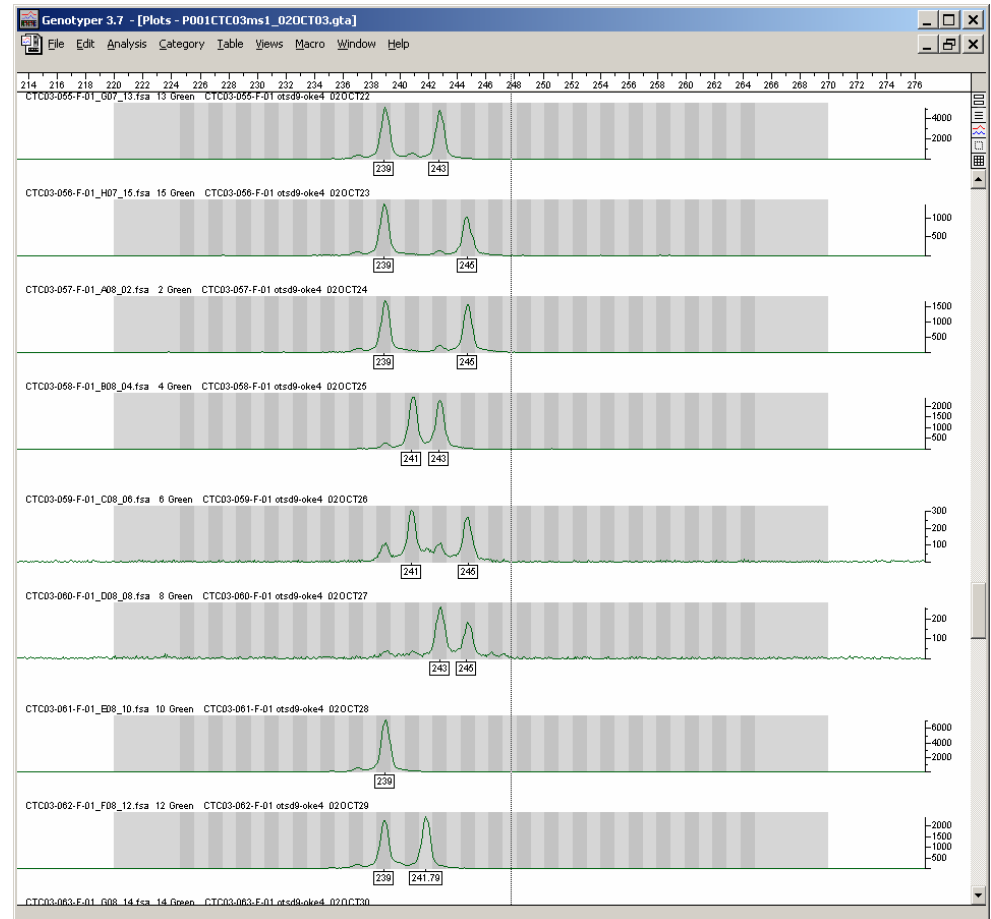
"C" — GTCGG**C**CACCCTG

Example genotypes

## SNP locus

Microsatellite locus

Typical GSI power function

- Power increases rapidly with increasing alleles

- The shape of the curve is similar for microsatellites and SNPs (ideal conditions)

- The number of alleles that obtains asymptotic power varies with their frequency and selection coefficient—not marker class

Power

Number of independent alleles
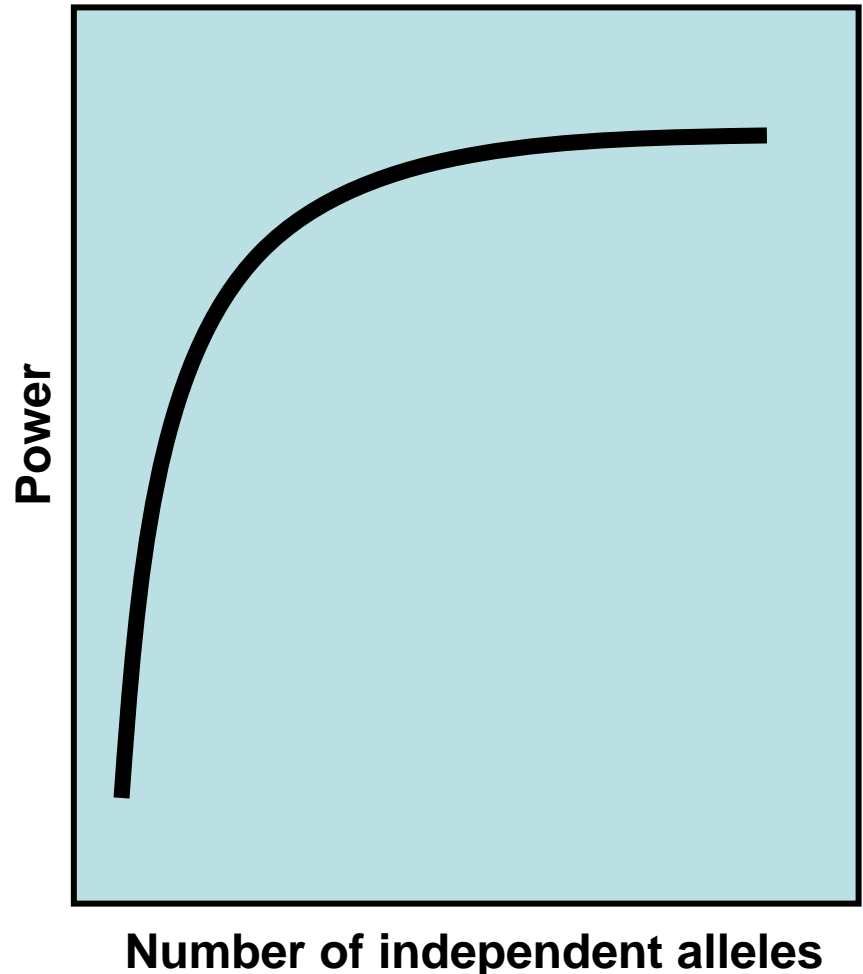
# Frequency distribution of GSI discriminatory power among alleles

•By chance, selection, or historical contingency, some alleles are highly informative, whereas others provide less

Power
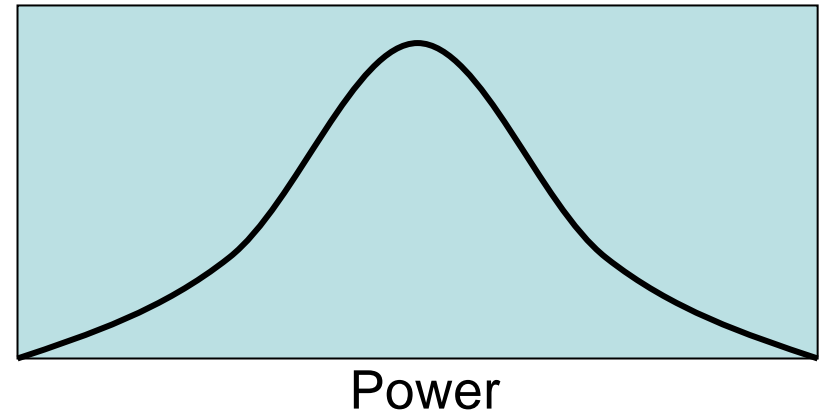
•Also by chance or other factors, some alleles show different patterns of relatedness among populations

Allele 1

Population A

Population B

Population C

Allele 2

Population A

Population B
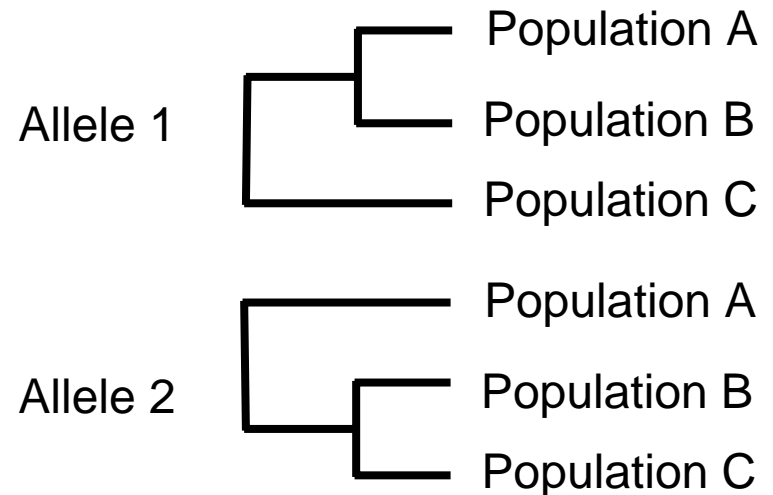
Population C

"Core" Populations
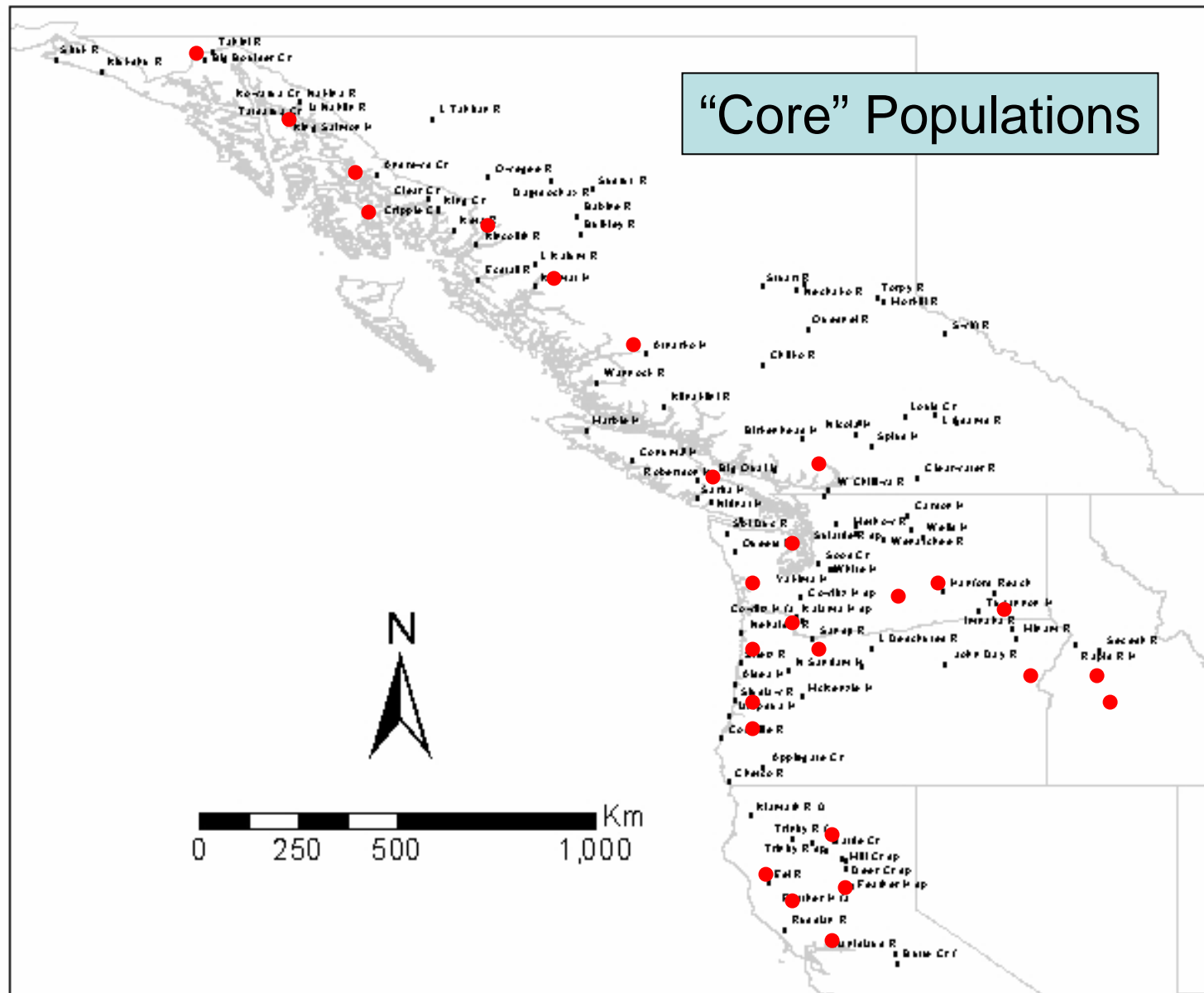
NWFSC's comment on the Expert Panel Report

"SNPs have some technical advantages over microsatellites, but the information they provide will not be fundamentally different or better than that what is currently available using microsatellites. To the degree that the PSC or other management agencies want to explore greater use of GSI methods for fishery management in the near future (e.g., within the next five years), we recommend that they take advantage of the considerable investment that has already been made in microsatellite technology."

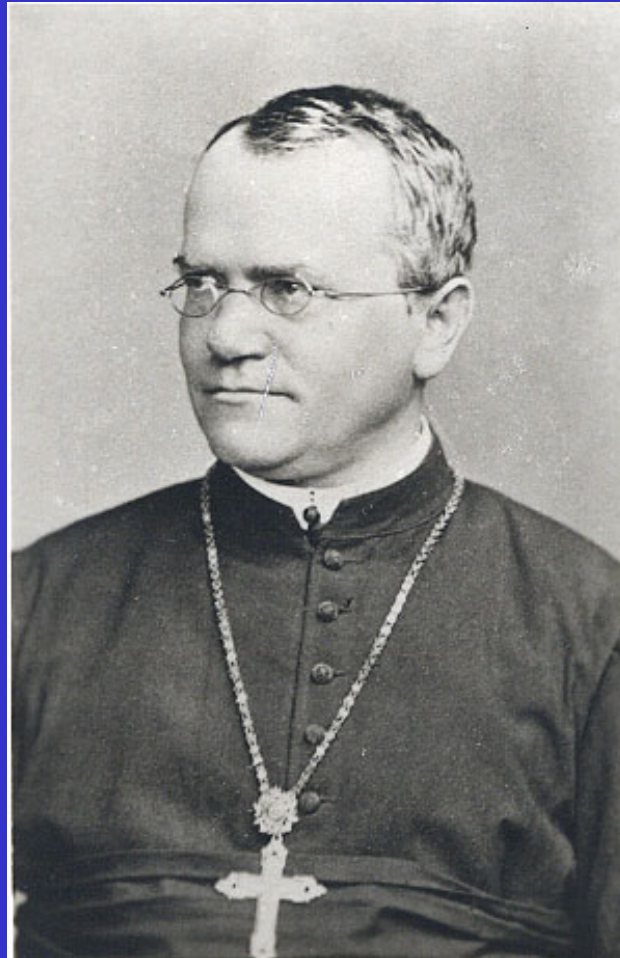# Full Parental Genotyping

GSI relies on temporally stable genetic differences among reporting groups—FPG does not

How FPG works

- All or nearly all potential parents are genotyped

- Samples of offspring are attributed to specific matings

- Routinely used for monitoring hatchery/wild interactions in semi-closed systems

"My time will come."

Gregor Johann Mendel 1822-1884

Mendelian inheritance
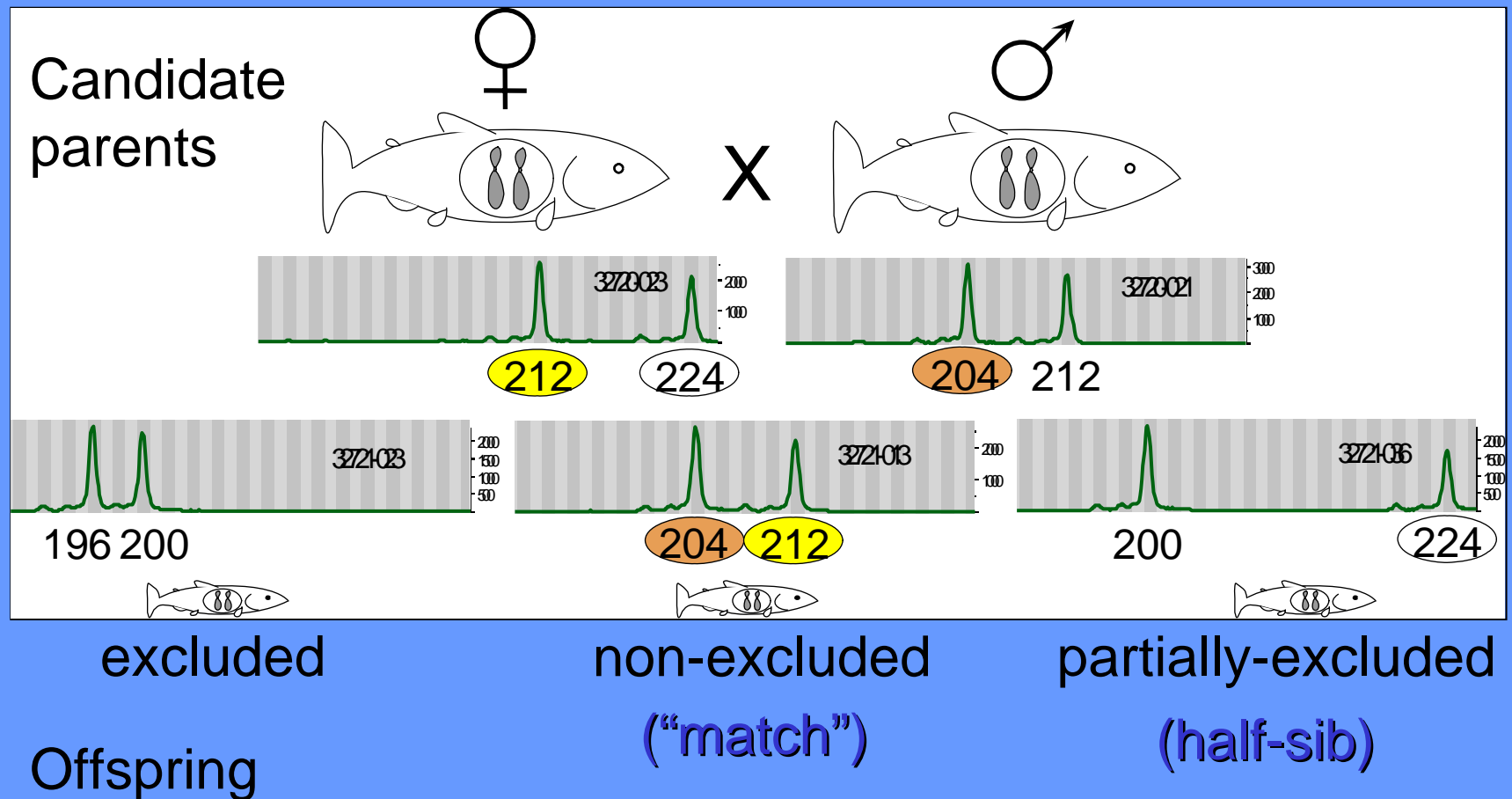
Steelhead returning to Little Sheep Creek

X

Trap intercepts
migrating adults,
March – May

Parent tissue
samples

Sample juvenile
offspring, Aug. – Oct.
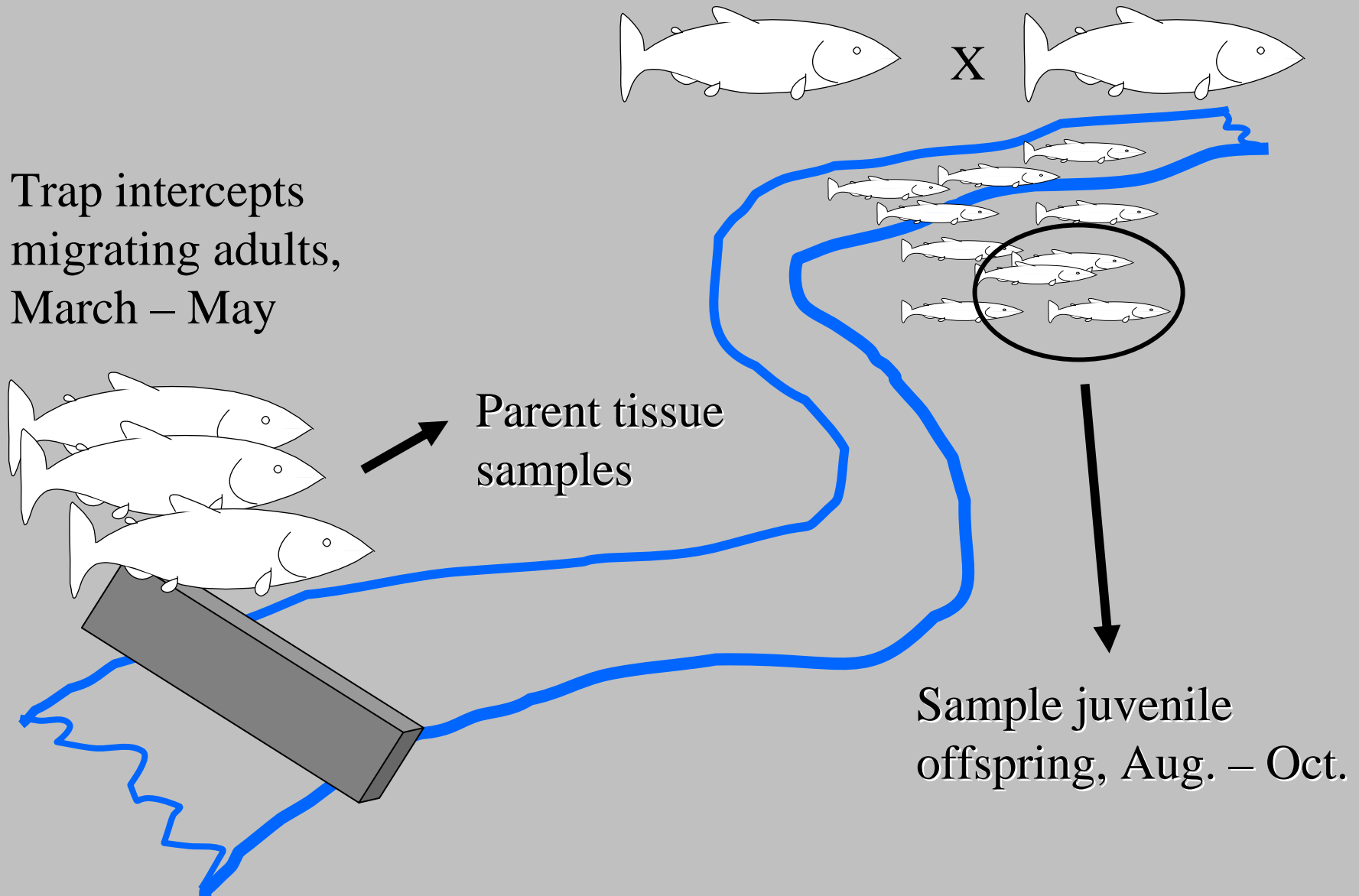
FPG caveats

- Unrealized exclusionary power typical of parentage studies
- Would not represent wild stocks
- Relies heavily on total marking
- Not necessarily integrated into GSI
- Cost

Genetic Analysis of Pacific Salmonids:

The GAPS consortium

GAPS Collaborators 2005

Anton Antonovich, Alaska Department of Fish and Game
Bill Ardren, US Fish and Wildlife Service, Abernathy
Melissa Baird, Northwest Fisheries Science Center
Michael Banks, Oregon State University
Terry Beacham, Canadian Department of Fisheries and Oceans
Reneé Bellinger, Oregon State University
Scott Blankenship, Southwest Fisheries Science Center
Craig Busak, Washington Department of Fish and Wildlife
Matt Campbell, Idaho Department of Fish and Game
Cheryl Dean, Southwest Fisheries Science Center
Nick Decovich, Alaska Department of Fish and Game
Carlos Garza, Southwest Fisheries Science Center
Chuck Guthrie, Alaska Fisheries Science Center, Auk Bay
Eric Kretschmer, US Fish and Wildlife Service, Anchorage
David Kuligowski, Northwest Fisheries Science Center
Tricia Lundrigan, University of Washington
Paul Moran, Northwest Fisheries Science Center
Shawn Narum, Columbia River Intertribal Fish Commission
Devon Perce, Southwest Fisheries Science Center
Madison Powell, University of Idaho
Lisa Seeb, Alaska Department of Fish and Game
Jim Shaklee, Washington Department of Fish and Wildlife
Christian Smith, Alaska Department of Fish and Game
Ora Stefanowski, US Fish and Wildlife Service
Jeff Stephenson, Columbia River Intertribal Fish Commission
Janine Supernault, Canadian Department of Fisheries and Oceans
David Teel, Northwest Fisheries Science Center
Bill Templin, Alaska Department of Fish and Game
Don VanDoornik, Northwest Fisheries Science Center
Ken Warheit, Washington Department of Fish and Wildlife
John Wenburg, US Fish and Wildlife Service
Richard Wilmot, Alaska Fisheries Science Center
Ruth Withler, Canadian Department of Fisheries and Oceans
Sewall Young, Washington Department of Fish and Wildlife

Challenges to microsatellite standardization

Few loci in common
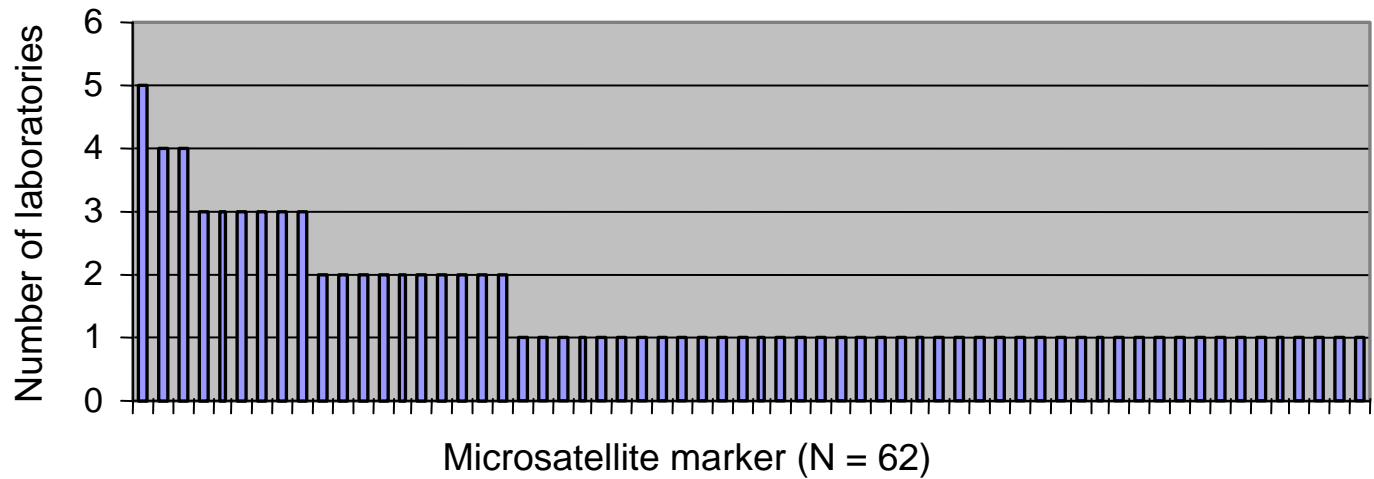
Precise but inaccurate sizing

Microsatellite marker (N = 62)

Fragment size (bp)
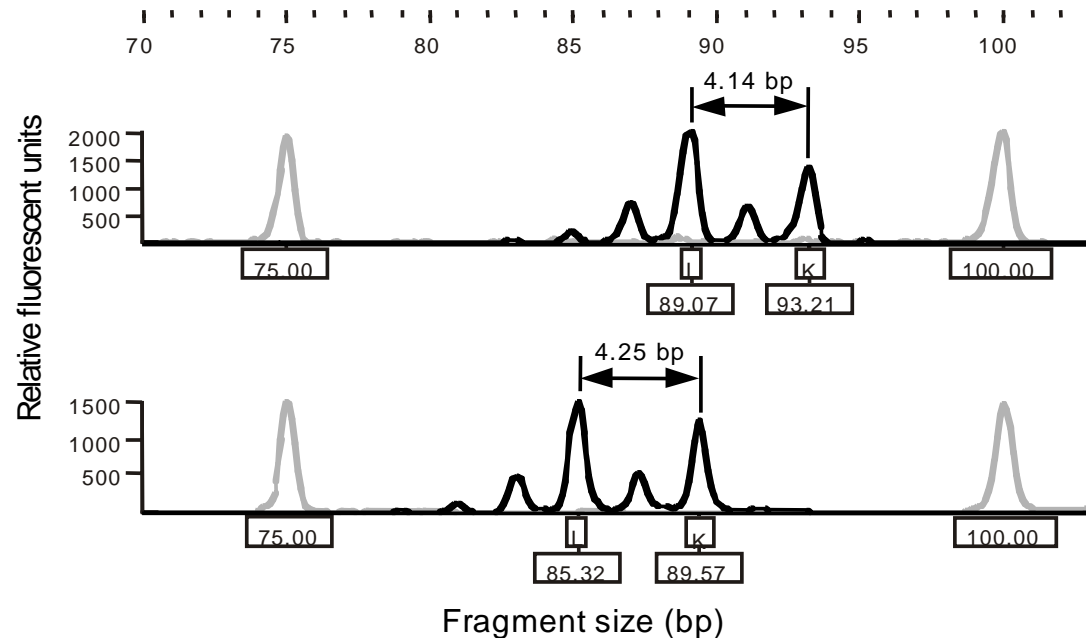
Chinook microsatellite standardization objectives

- Develop a standardized coast-wide DNA baseline that will resolve lineages and stocks within the region for which the CTC is responsible

- Facilitate future expansion of the DNA baseline and addition of novel genetic markers through exchange of DNA samples

- **Develop a database application to support the dissemination and growth of the baseline**

Partially corrected concordance

B) Partially corrected

Locus	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Lab 8	Lab 9	Average
Ogo2	0.987	1.000	1.000	1.000	0.993	0.988	1.000	0.993	1.000	0.996
Ogo4	0.994	1.000	1.000	0.995	1.000	0.968	0.995	0.994	0.990	0.993
Oki100	0.978	1.000	1.000	1.000	1.000	1.000	1.000	0.970	1.000	0.994
OMM1080	1.000	1.000	0.995	1.000	1.000	0.938	1.000	0.994	1.000	0.995
Ots201b	0.984	1.000	1.000	1.000	1.000	0.993	0.995	0.985	1.000	0.990
Ots208b	0.994	1.000	1.000	1.000	1.000	1.000	0.995	0.970	0.995	0.996
Ots211	1.000	1.000	0.994	1.000	0.993	0.955	0.994	0.985	0.994	0.994
Ots212	0.989	1.000	1.000	1.000	1.000	0.989	0.995	0.994	1.000	0.992
Ots213	0.987	1.000	0.982	1.000	0.985	0.994	1.000	1.000	1.000	0.999
Ots3M	1.000	1.000	0.988	0.994	1.000	0.949	1.000	1.000	0.995	0.980
Ots9	1.000	1.000	1.000	1.000	1.000	0.979	1.000	1.000	1.000	0.991
OtsG474	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
Ssa408	0.987	0.929	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.991
Average	0.992	0.995	0.997	0.999	0.998	0.981	0.998	0.991	0.998	0.994
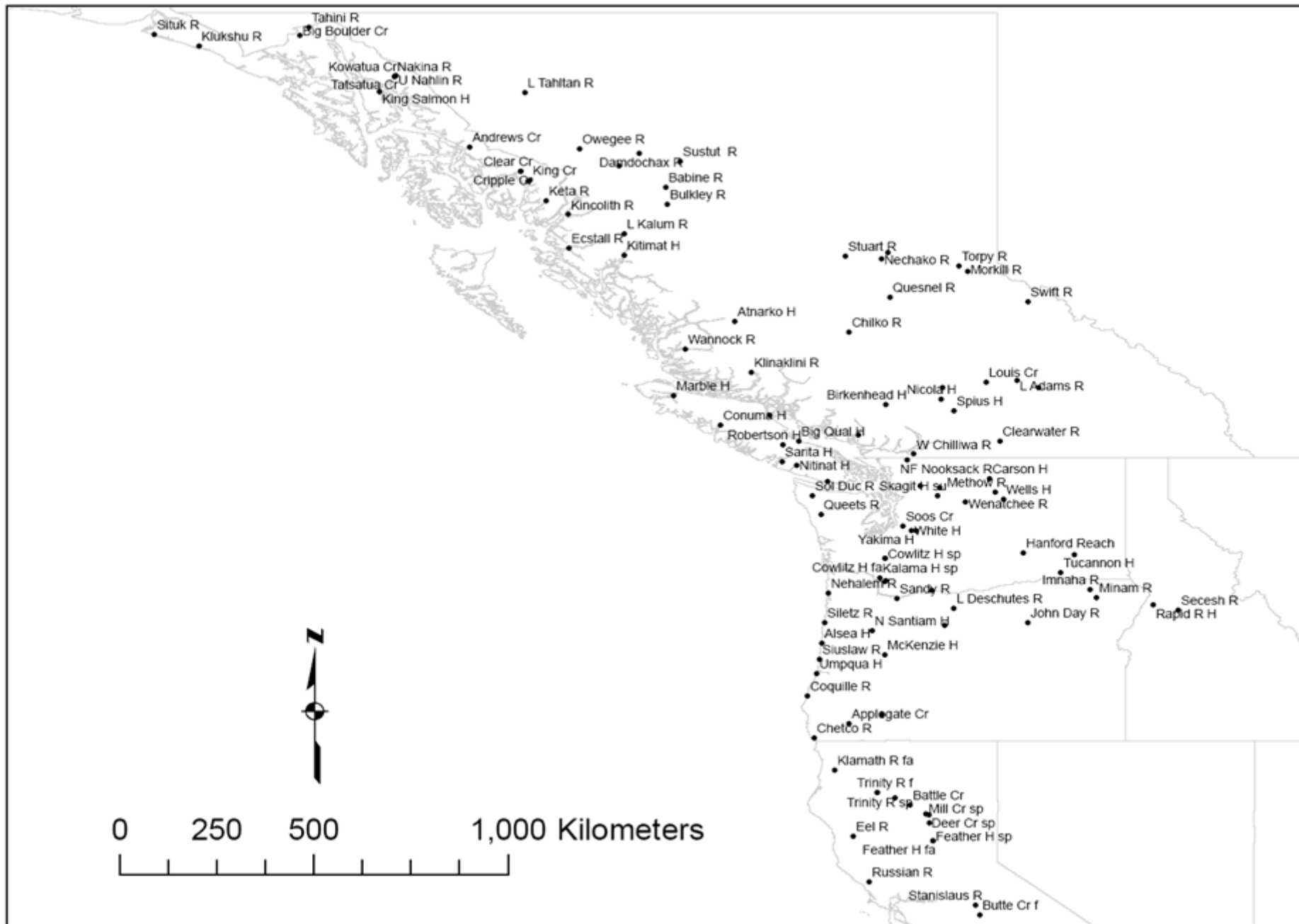
Current Chinook coast-wide baseline

- ~22,300 individual fish genotyped
- 166 populations (>300 samples)
- 42 regions, Gulf of AK to Central Valley CA
- Preliminary analyses complete
  - Descriptive population genetics, phylogeography
  - GSI power analysis (Seeb et al. *accepted*)

Issues raised in peer review

- How will GAPS assure data access to the research community in the future?

- How will future labs become standardized who did not participate in the original effort?*

* http://www.nwfsc.noaa.gov/research/divisions/cbd/documents/gaps_year2_final.pdf

Issues raised in peer review

Two PSC-funded projects are relevant

- How will GAPS assure data access to the research community in the future?

The GAPS database

- How will future labs become standardized who did not participate in the original effort?

Allele ladders and database documentation

GAPS collaborators
PSC database project

Auke Bay Laboratory, NOAA
Gene Conservation Laboratory, ADFG
Pacific Biological Station, CDFO
Northwest Fisheries Science Center, NOAA
Southwest Fisheries Science Center, NOAA
Washington Department of Fish and Wildlife

Goals of the database project

To provide a safe and secure repository for standardized, shared, interagency, genetic data—Chinook salmon *et al.*

GAPS database elements

- Security—access and authentication
- Interactive *ad hoc* query tools
- Multiple download formats and custom reports
- Interactive graphics
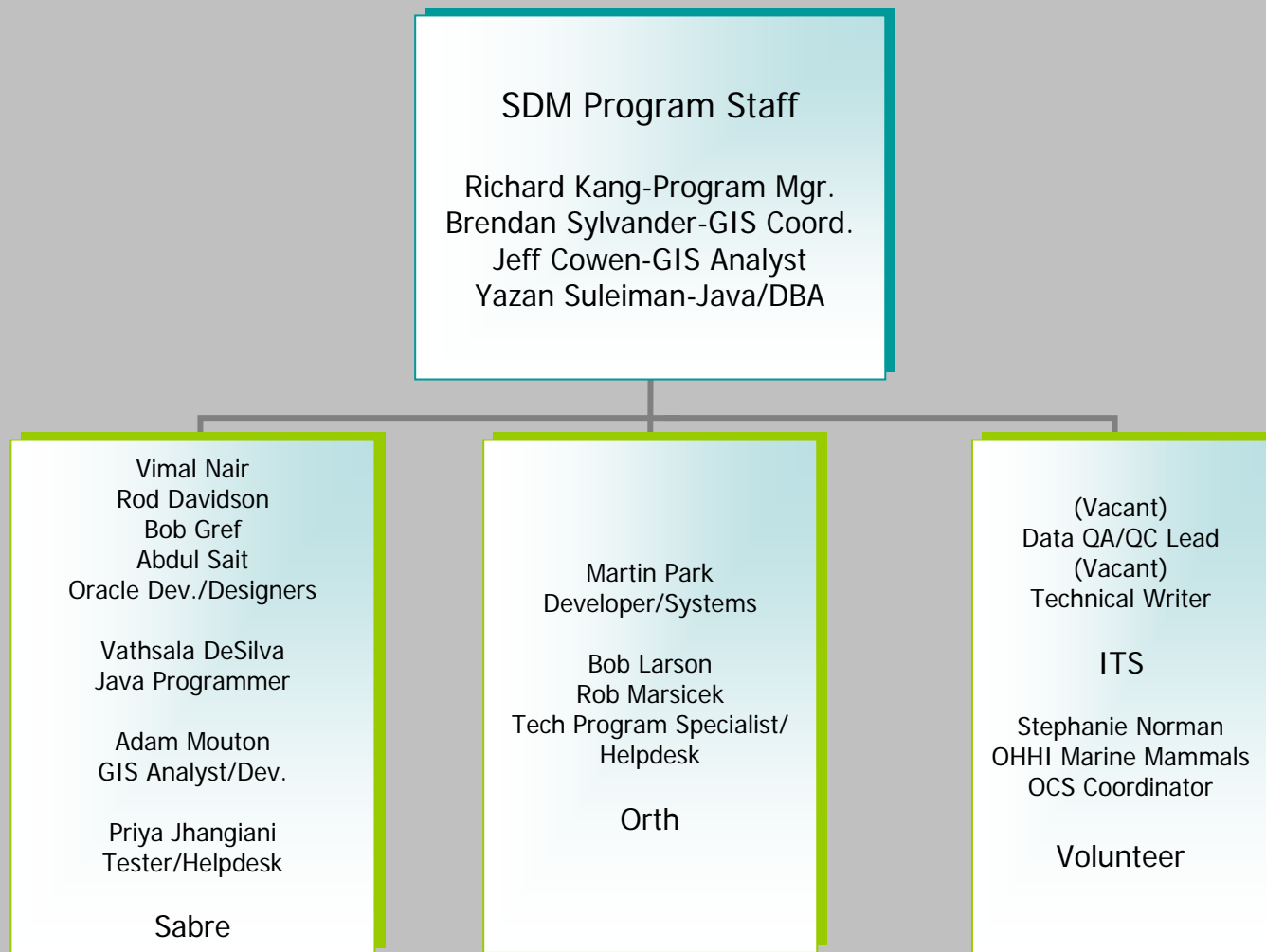- Web-enabled mapping tools

Database features

- Centralized, secure, web-based application
- Business rules for data submission, QC
- Flexible and dynamic
- Historical continuity, documentation of changes
- Transportable

Northwest Fisheries Science Center
Scientific Data Management
4 FTE's, 10 Contractors & 1 Volunteer

SDM Program Staff

Richard Kang-Program Mgr.
Brendan Sylvander-GIS Coord.
Jeff Cowen-GIS Analyst
Yazan Suleiman-Java/DBA

Vimal Nair
Rod Davidson
Bob Gref
Abdul Sait
Oracle Dev./Designers

Vathsala DeSilva
Java Programmer

Adam Mouton
GIS Analyst/Dev.

Priya Jhangiani
Tester/Helpdesk

Sabre

Martin Park
Developer/Systems

Bob Larson
Rob Marsicek
Tech Program Specialist/
Helpdesk

Orth

(Vacant)
Data QA/QC Lead
(Vacant)
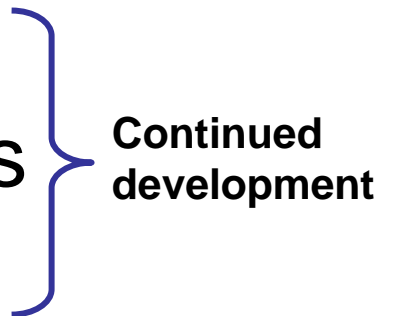Technical Writer

ITS

Stephanie Norman
OHHI Marine Mammals
OCS Coordinator

Volunteer

Northwest Fisheries Science Center
Scientific Data Management
Hardware and Software Infrastructure

Current status

✓ • IT and genetics 101, dictionaries

✓ • Information collection, tables & relationships

✓ • Schema design

✓ • Population of database

✓ • Development of query structure

✓ • Output formats and custom reports

✓ • Graphics and mapping

Continued development

✓ • Support for inter-lab standardization

File   Edit   View   Favorites   Tools   Help

Back | Search | Favorites
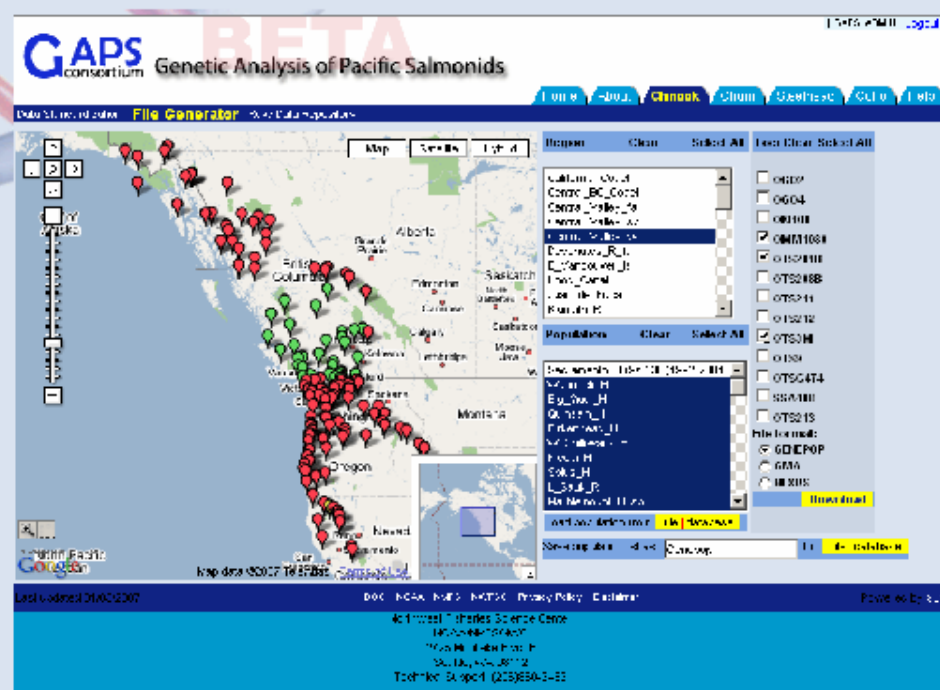
Address http://devedb.nwfsc.noaa.gov/pls/htmldb/f?p=113:51:13185955714432880960::::::

| PAUL.MORAN | Change Password | Logout |

GAPS consortium Genetic Analysis of Pacific Salmonids

BETA

Home | About | **Chinook** | Chum | Steelhead | Coho | Help

| **Data Standardization** | File Generator | Raw Data Repository |

Data Summary / **Genotypic Data** / Sanctioned Alleles

Genotypic Data

[Reset]

Search

[   ] [Search]

Region Selector(# of Individuals) [help]

| All |
| California_Coast (281) |
| Central_BC_Coast (429) |
| Central_Valley_fa (432) |

[Go >>]

Population(# of Individuals) [help]

| All |
| California_Coast: Eel_R(137) |
| California_Coast: Russian_R(144) |
| Central_BC_Coast: Atnarko_H(144) |

[Go]

row(s) 1 - 15 of 22411   Next ▶

REGION	POPULATION	RUN	RUN2	SAMPLE	ORIGIN ▲	STAGE	COLLECTOR	COLLECTION DATE	COLLECTION METHOD	INDIVIDUAL	GENOTYPER
Central_BC_Coast	Kitimat_H	F	Fa	Kitimat_H_97	H	Adult	CDFO	1997		2700	CDFO
Central_BC_Coast	Wannock_H	F	Fa	Wannock_H_96	H	Adult	CDFO	1996		3500	CDFO
Central_BC_Coast	Atnarko_H	Fa	Fa	Atnarko_H_96	H	Adult	CDFO	1996		3900	CDFO
Central_BC_Coast	Kitimat_H	F	Fa	Kitimat_H_97	H	Adult	CDFO	1997		2608	CDFO
Central_BC_Coast	Wannock_H	F	Fa	Wannock_H_96	H	Adult	CDFO	1996		3600	CDFO
Central_BC_Coast	Kitimat_H	F	Fa	Kitimat_H_97	H	Adult	CDFO	1997		2609	CDFO
Central_BC_Coast	Kitimat_H	F	Fa	Kitimat_H_97	H	Adult	CDFO	1997		2611	CDFO
Central_BC_Coast	Kitimat_H	F	Fa	Kitimat_H_97	H	Adult	CDFO	1997		2613	CDFO

Last updated 05/14/2007

DOC | NOAA | NMFS | NWFSC | Privacy Policy | Disclaimer

Powered by SDM

Northwest Fisheries Science Center
NOAA/NMFS/NWC
2725 Montlake Blvd. E

Internet

| PAUL.MORAN | Change_Password | Logout |

BETA

GAPS consortium Genetic Analysis of Pacific Salmonids

Home | About | **Chinook** | Chum | Steelhead | Coho | Help

| Data Standardization | **File Generator** | Raw Data Repository |

Map | Satellite | **Hybrid**
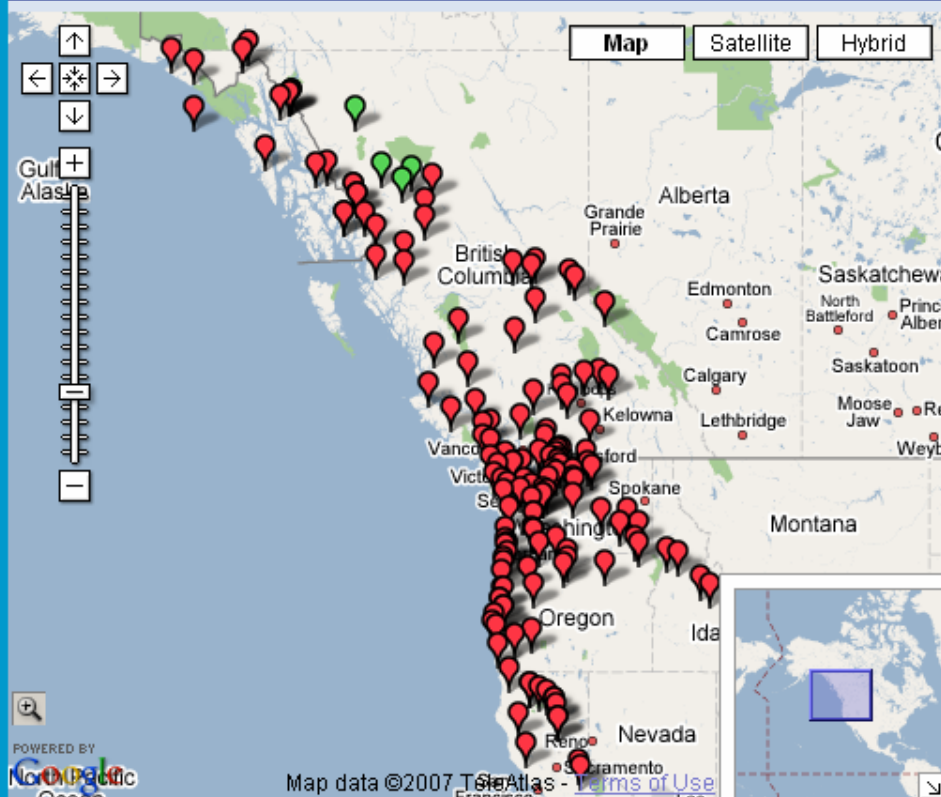
Region   **Clear Region**   **Select All**

California_Coast
Central_BC_Coast
Central_Valley_fa
Central_Valley_sp
Central_Valley_wi
Deschutes_R_fa
E_Vancouver_Is
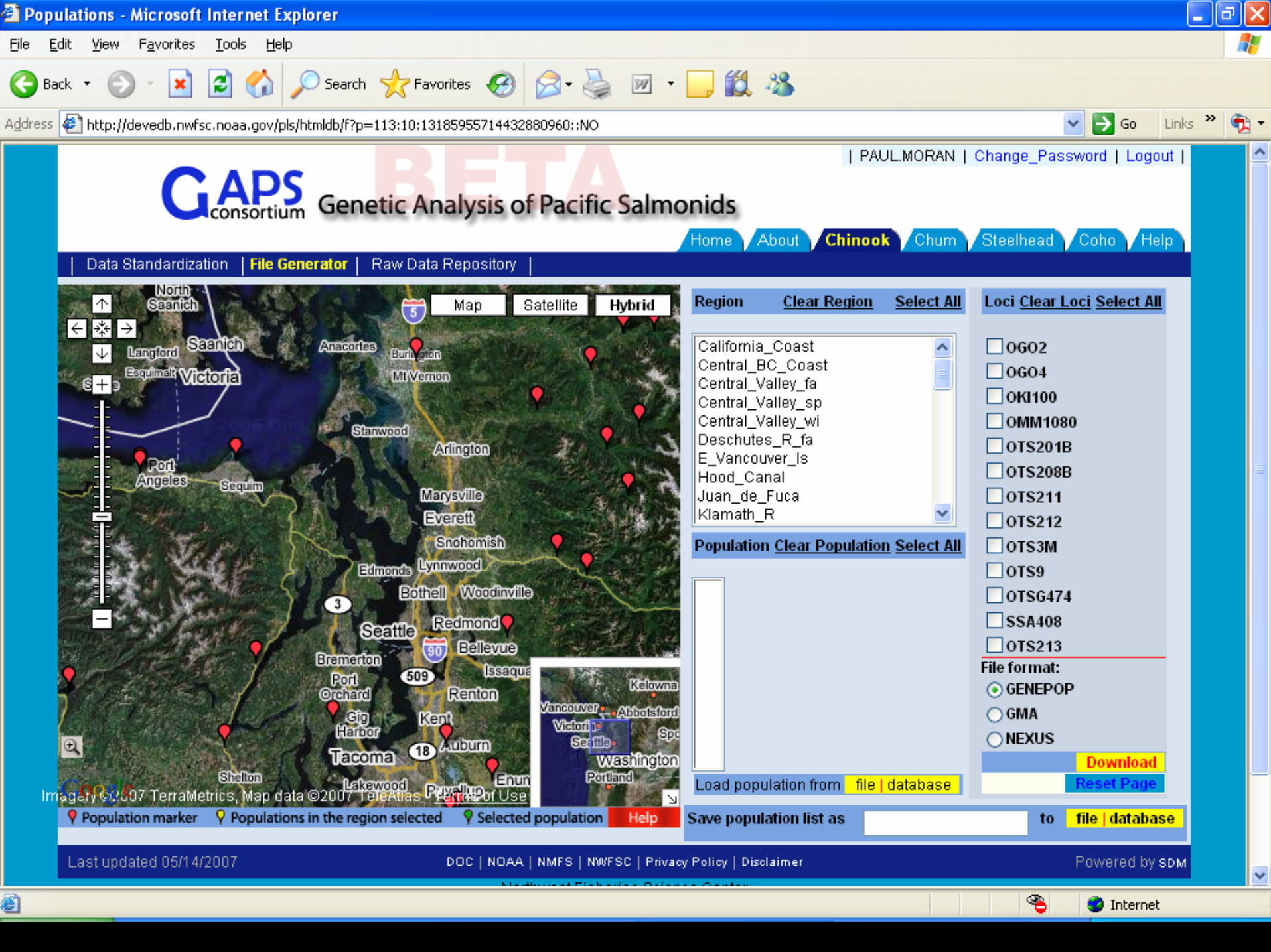Hood_Canal
Juan_de_Fuca
Klamath_R

Population Clear Population Select All

Load population from file | database

Loci Clear Loci Select All

☐ OGO2
☐ OGO4
☐ OKI100
☐ OMM1080
☐ OTS201B
☐ OTS208B
☐ OTS211
☐ OTS212
☐ OTS3M
☐ OTS9
☐ OTSG474
☐ SSA408
☐ OTS213

File format:
◉ GENEPOP
◯ GMA
◯ NEXUS

Download
Reset Page

♥ Population marker   ♥ Populations in the region selected   ♥ Selected population   Help   Save population list as   to file | database

Last updated 05/14/2007   DOC | NOAA | NMFS | NWFSC | Privacy Policy | Disclaimer   Powered by SDM

Imagery ©2007 TerraMetrics, Map data ©2007 TeleAtlas

Internet

| PAUL.MORAN | Change Password | Logout |

BETA

GAPS consortium Genetic Analysis of Pacific Salmonids

Home | About | **Chinook** | Chum | Steelhead | Coho | Help

| **Data Standardization** | File Generator | Raw Data Repository |

Data Summary | Genotypic Data | **Sanctioned Alleles**

Sanctioned Alleles

Allele	Category	Holotype	Description	Notes
200	Ogo2	BL-ADFG-KANDR89_12	Ogo2 curator doc 4-17-06	
202	Ogo2	CTC03-018	Ogo2 curator doc 4-17-06	
206	Ogo2	CTC03-316	Ogo2 curator doc 4-17-06	
208	Ogo2	BL-CRITFC-B23-49	Ogo2 curator doc 4-17-06	
210	Ogo2	BL-ADFG-KKOW89_16	Ogo2 curator doc 4-17-06	
212	Ogo2	CTC03-131	Ogo2 curator doc 4-17-06	
214	Ogo2	CTC03-366	Ogo2 curator doc 4-17-06	
216	Ogo2	CTC03-188	Ogo2 curator doc 4-17-06	
218	Ogo2	CTC03-023	Ogo2 curator doc 4-17-06	
220	Ogo2	CTC03-186	Ogo2 curator doc 4-17-06	
222	Ogo2	CTC03-137	Ogo2 curator doc 4-17-06	
224	Ogo2	CTC03-186	Ogo2 curator doc 4-17-06	
226	Ogo2	CTC03-313	Ogo2 curator doc 4-17-06	
228	Ogo2	CTC03-444	Ogo2 curator doc 4-17-06	
230	Ogo2	CTC03-013	Ogo2 curator doc 4-17-06	
232	Ogo2	CTC03-364	Ogo2 curator doc 4-17-06	
234	Ogo2	CTC03-235	Ogo2 curator doc 4-17-06	
236	Ogo2	CTC03-318	Ogo2 curator doc 4-17-06	

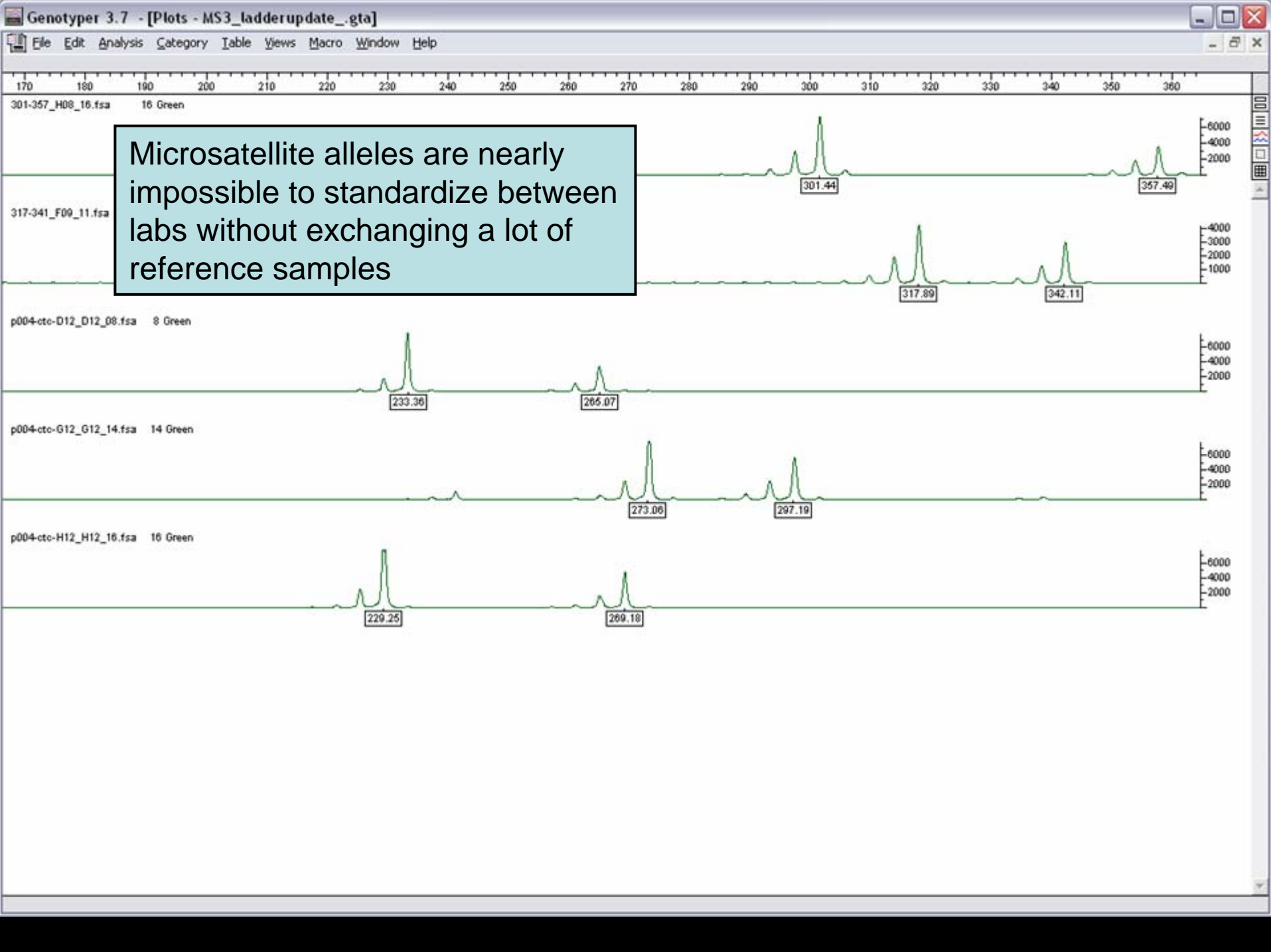Last updated 05/14/2007   DOC | NOAA | NMFS | NWFSC | Privacy Policy | Disclaimer   Powered by SDM
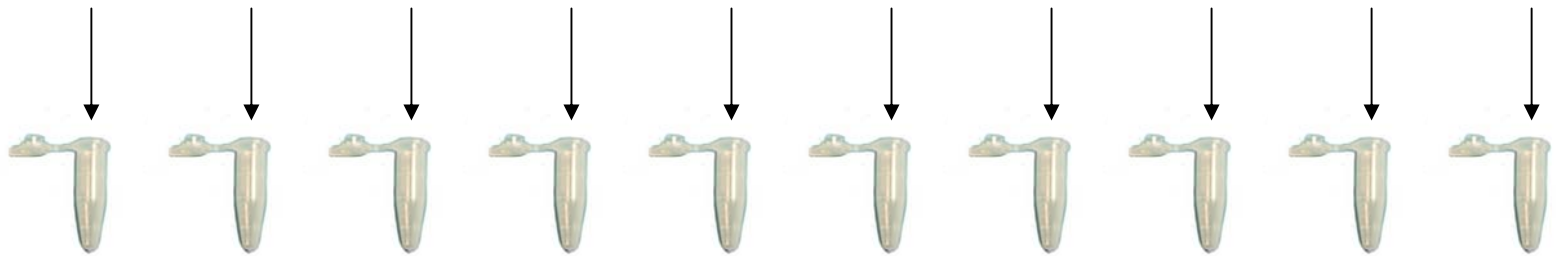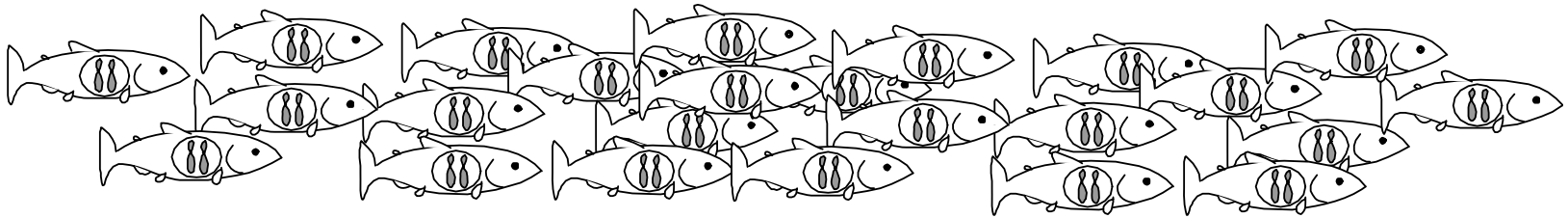
Northwest Fisheries Science Center
NOAA/NMFS/NWC
2725 Montlake Blvd. E

Future of standardized data

- Allele ladders will dramatically simplify microsatellite standardization and certification of new labs

- SNP data will be readily incorporated with required documentation for standardization

- Implementation of standardized databases for chum, coho, sockeye, and steelhead

- Repository for old and new data, fully and partially standardized, etc.

Microsatellite alleles are nearly impossible to standardize between labs without exchanging a lot of reference samples

What is an Allele Ladder?

Combined microsatellite products

An allele ladder provides all the reference "samples" in a single assay

Rungs on the ladder align with "unknown" alleles providing robust inter-laboratory standardization

Please see LaHood et al. 2002 (Molecular Ecology Notes **2**:187–190) for use of ladders in inter-laboratory standardization

Future of standardized data

- Allele ladders will dramatically simplify microsatellite standardization and certification of new labs

- SNP data will be readily incorporated with required documentation for standardization

- Implementation of standardized databases for chum, coho, sockeye, and steelhead

- Repository for old and new data, fully and partially standardized, etc.
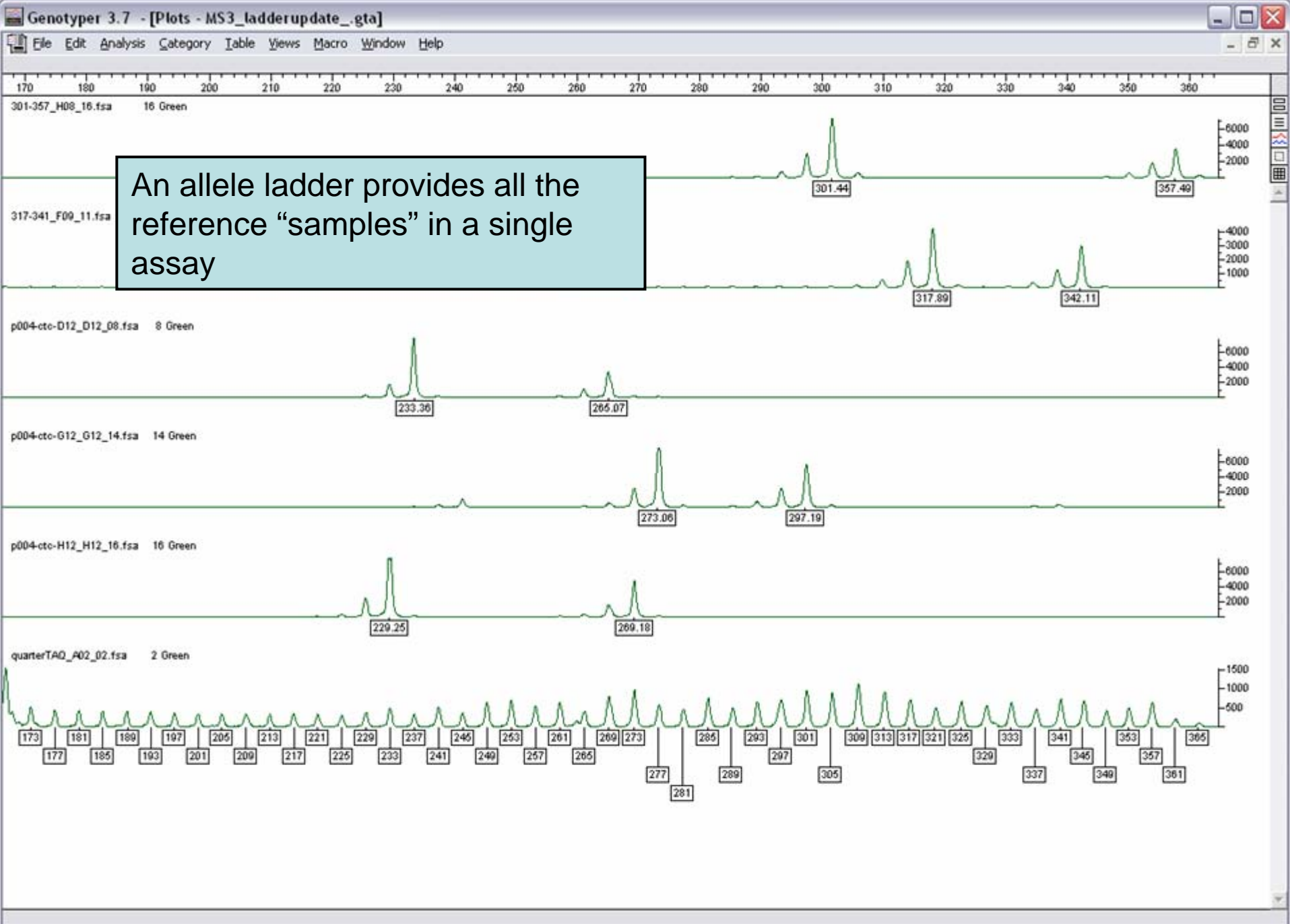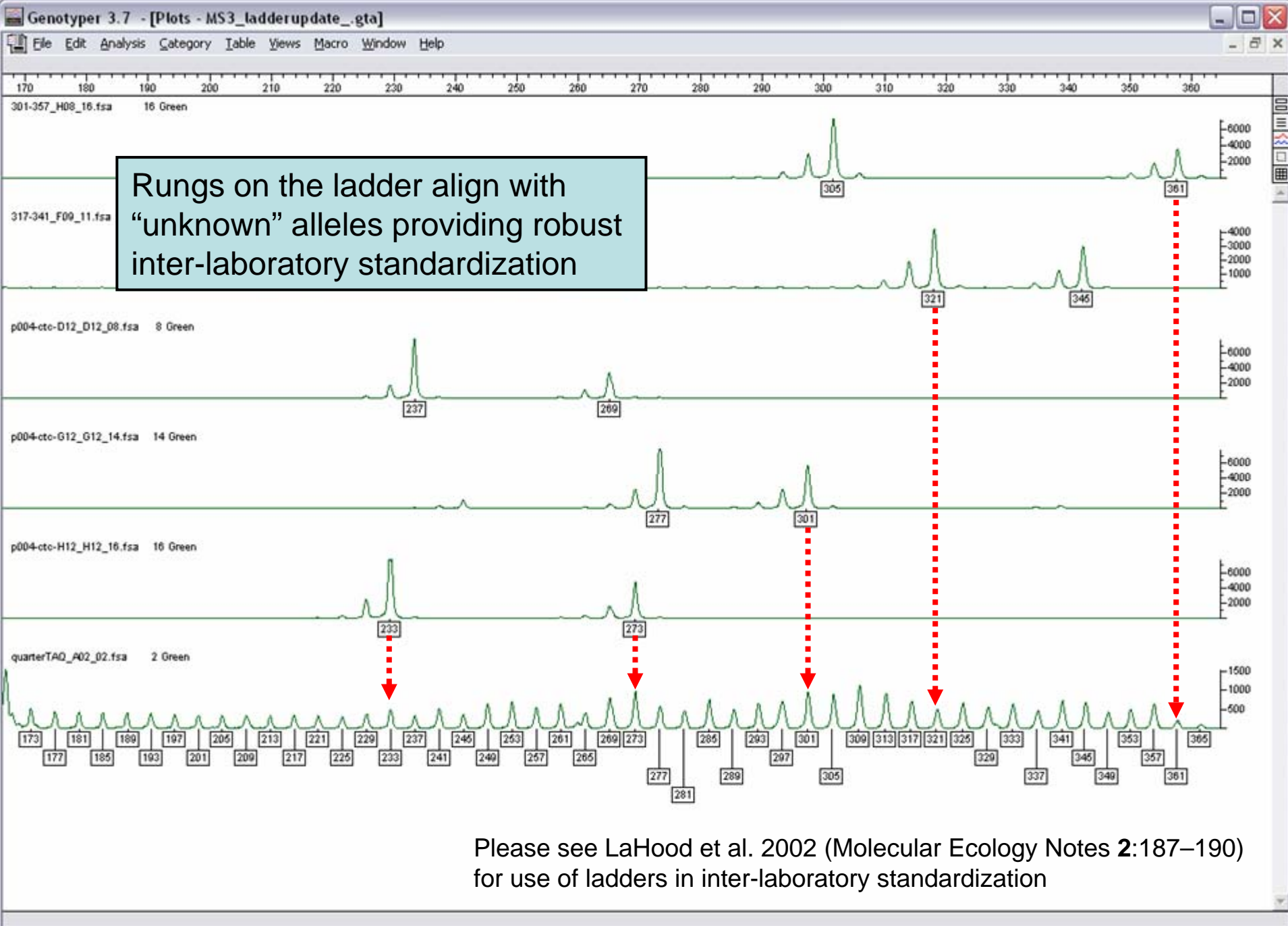
Current opportunities and limitations for genetics in fishery management

Current opportunities

- Realtime GSI for terminal fishery mgmt—DFO, John Candy
- Ocean harvest impacts on depressed wild stocks— Klamath R test fisheries
- Sub-legal encounter rates—SEAK troll
- Ecological genetic applications
  - CROOS ocean ecology—individual assignment
  - Migration and life history variation—plume studies, estuarine habitat use
  - Stock-specific disease and parasite prevalence, ecotoxicology, (eg Rice et al poster)
  - Adaptive divergence among regions/life history types—EST-linked SNPs and microsatellites
  - FPG—could be done coastwide—right now for <$10M for "tagging"

Current challenges

- Genetic issues—marker classes, mixture analysis methods

- IT issues—creating a permanent, safe, secure and convenient web application (funding)

- Fishery management issues—how, when, and where to implement genetic methods (if at all)

The other side of controversy--beyond marker wars to World peace

- Potential missed opportunities
- The roots of controversy
- DNA problems and potential mitigations
- Shared view of CWT augmentation

("I have a dream")

Potential missed opportunities

- Less efficient use of limited mgmt funds for improved encounter rate estimates, wild encounter rates, better proxies
- No additional classes of potentially useful information to harvest—direct or indirect (ecological genetics, basic research)
- Lost expertise—some of the pioneers of West Coast GSI, have moved on to other problems
- Chinook salmon and the fishery resource suffer if arguments over methods prevent progress

The roots of controversy

- Different agency perspectives—high seas versus terminal fisheries, continent of origin versus hatchery/wild impacts, regional focus
- Different professional judgments—timing for implementation of SNPs, scale and specific application of DNA methods, cost and efficiency projections for DNA and modified CWT programs
- Different "schools"—CWTs, PIT tags, radiotags, microchem & genetics
- Historical trajectory—Treaty constraints (everything above is historical, but this is a current show stopper)

Problems with DNA

- No age structure from conventional GSI
- Probabilistic, rather than deterministic, stock of origin (IA)
- Poor estimates at low encounter rates
- Lack of agreement on methods
- Excessive cost

Potential solutions or mitigations

- Scale aging might be possible if stratified by stock (at a scale of ~50 stocks coast-wide)
- No method is without error—modern GSI methods are powerful, but have practical limits
- Every fish is marked genetically—power only related to scale, which has increased exponentially with declining costs
- Increasing consensus on value of genetic data, yet may be fundamentally unsuited for managing treaty fisheries—guidance from managers

Developing a common vision for the appropriate implementation of DNA methods

- Interagency standardization is essential—arguments about methods confound management debate
- Balance independence versus standardization, unified versus discrete funding
- Create a product or a system that is more useful than a single agency can generate alone
  - Facilitate many contributors, robust web app with Genbank-like data submission filters, data submission contract requirements of funders, and allele ladders
  - Frequent updates, a catch-22 proposition—users won't update a database that is not useful or current
- Focus on biology rather than specific fisheries
- Play to strengths of a given method—the formula for 30 years of CWT success