

# **DNA-based stock-of-origin assignments of Chinook salmon smolts emigrating past Chandler Trap in 2004**

Todd W. Kassler, WDFW Genetics Laboratory  
Mark Johnston, Yakama Nation Fisheries  
James B. Shaklee, WDFW Genetics Laboratory  
Sewall F. Young, WDFW Genetics Laboratory

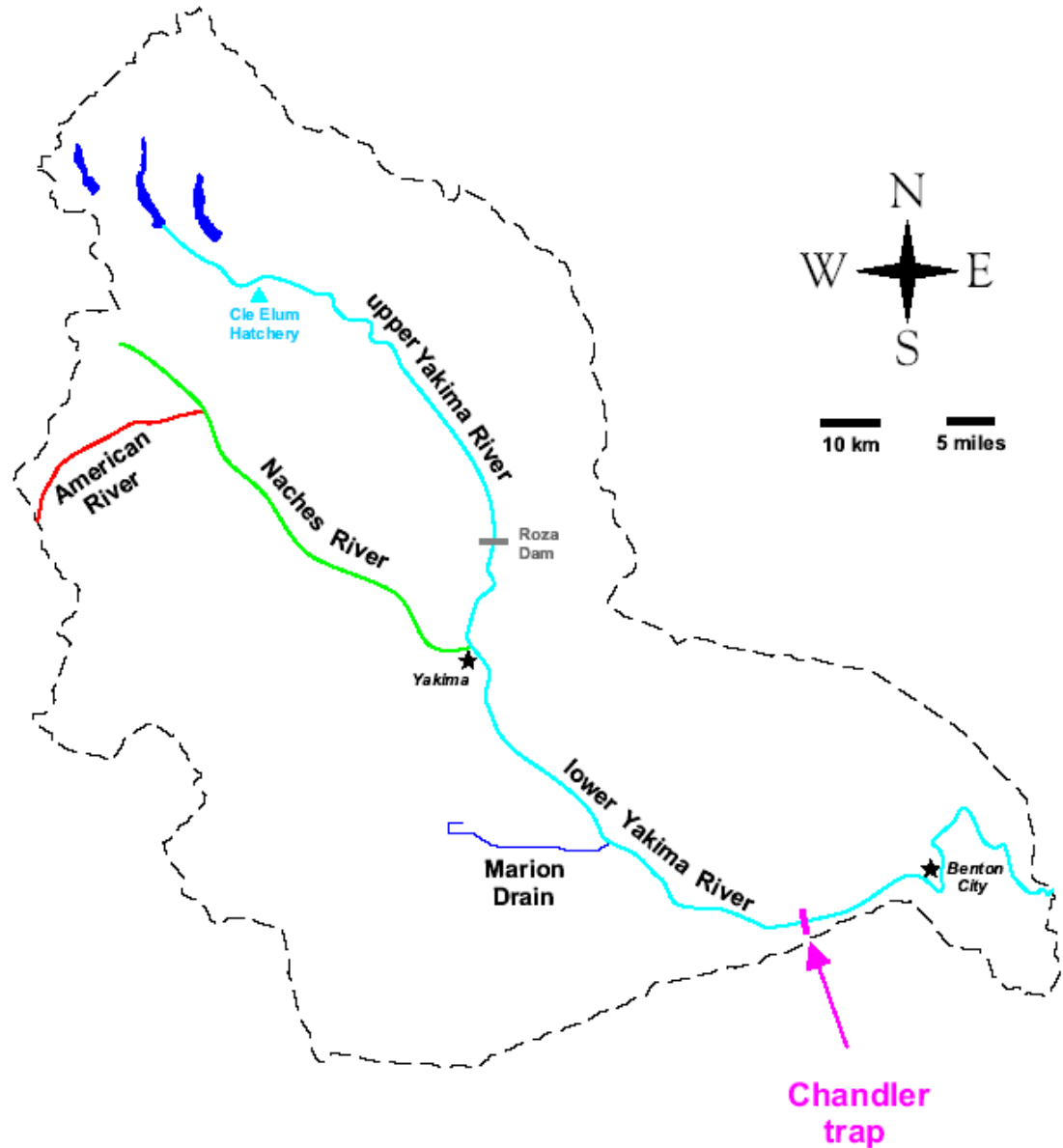
# Yakima basin Chinook salmon populations

spring Chinook populations:

- upper Yakima River
- American River
- Naches River

fall Chinook populations:

- lower Yakima River
- Marion Drain



# Primary objective

- To estimate the productivity of the spring Chinook salmon stocks in the Yakima basin
  - part of monitoring and evaluating Cle Elum Hatchery program
  - emigrating smolts are sampled at the Chandler trap
  - spring chinook smolts from American, Naches, and upper Yakima populations are comingled
  - a portion of the lower Yakima fall chinook population spawns above Chandler and must migrate past Chandler trap

# Ancillary objective

- To estimate other stock-specific biological / ecological parameters
  - migration timing
  - condition factor
  - disease incidence
  - other

# Genetics-based mixture analyses

- Mixed stock analyses (MSA) – use “baseline” data to estimate stock proportions in a mixture
- Clustering methods – attempt to identify genotypic clusters (= stocks) in a mixture without assuming knowledge of stock structure
- Individual assignments – use “baseline” data to assign individuals to stock-of-origin; stock proportions estimated from tallies

# The process

- Identify discrete populations and collect baseline samples

# The process

- Identify discrete populations and collect baseline samples
- **Genotype baseline samples**

# The process

- Identify discrete populations and collect baseline samples
- Genotype baseline samples
- **Sample the mixture**



# The process

- Identify discrete populations and collect baseline samples
- Genotype baseline samples
- Sample the mixture
- **Genotype mixture samples**

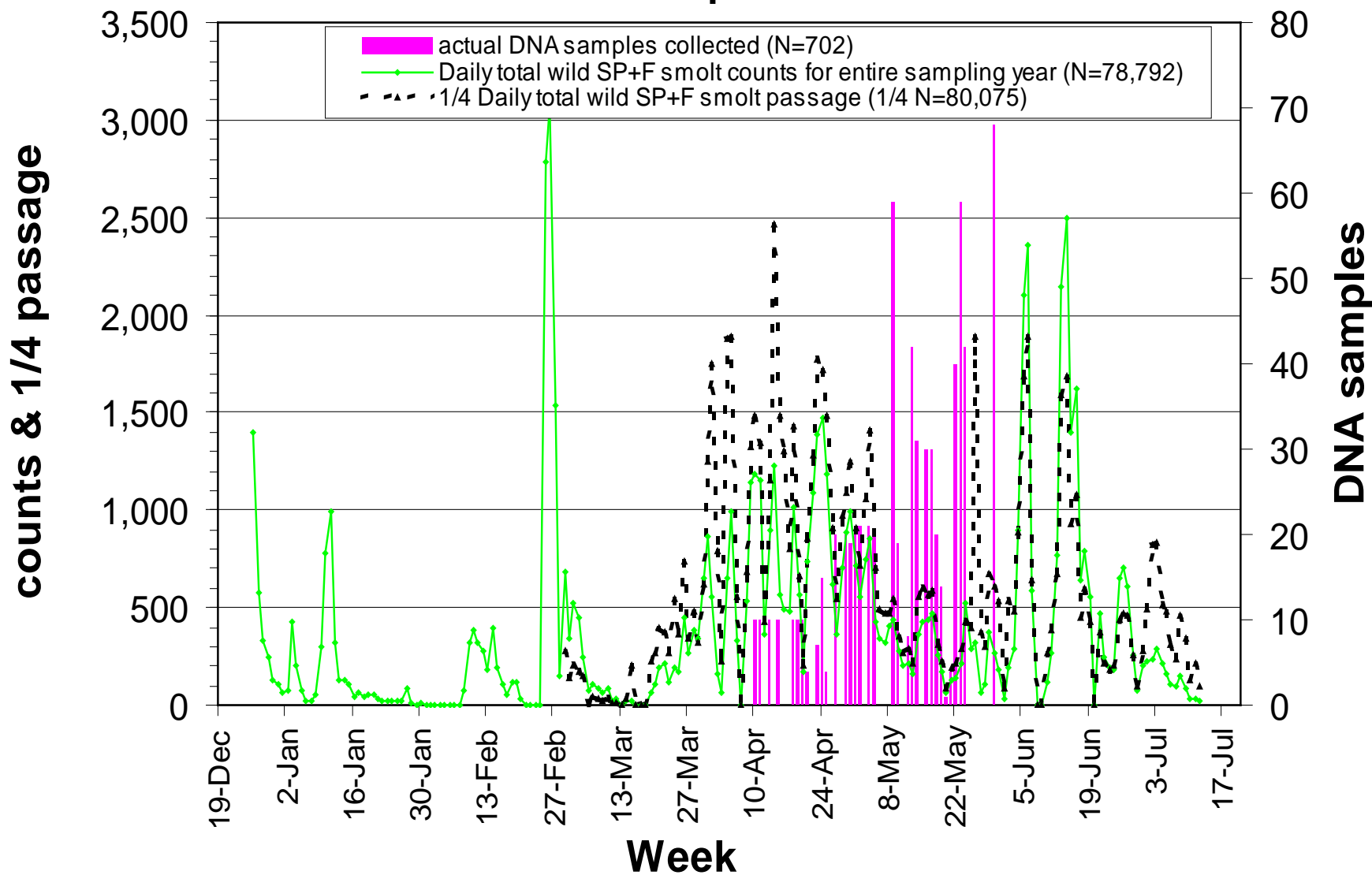
# The process

- Identify discrete populations and collect baseline samples
- Genotype baseline samples
- Sample the mixture
- Genotype mixture samples
- Perform stock-of-origin assignments using observed allele frequencies in baselines as expected frequencies in populations

# The quality of the analysis depends on representative sampling

- Identify discrete populations and collect baseline **samples**
- Genotype baseline **samples**
- **Sample** the mixture
- Genotype mixture **samples**
- Perform stock-of-origin assignments using observed allele frequencies in baselines as expected frequencies in populations

# Chandler trap 2003: counts, passage estimates, and genetic samples



# Chandler Trap - wild SP+F chinook counts and percentages

	2000		2001		2002		2003		average
	N	%	N	%	N	%	N	%	%
December	4	0.0%	2,032	1.9%	11,959	9.9%	2,796	3.5%	3.8%
January	2,957	5.5%	1,159	1.1%	21,666	17.9%	4,207	5.3%	7.5%
February	304	0.6%	1,386	1.3%	29,323	24.2%	10,138	12.9%	9.7%
<b>"winter" (Dec-Feb)</b>	<b>3,265</b>	<b>6.1%</b>	<b>4,577</b>	<b>4.2%</b>	<b>62,948</b>	<b>52.0%</b>	<b>17,141</b>	<b>21.8%</b>	<b>21.0%</b>
<b>March</b>	<b>83</b>	<b>0.2%</b>	<b>3,163</b>	<b>2.9%</b>	<b>7,392</b>	<b>6.1%</b>	<b>6,771</b>	<b>8.6%</b>	<b>4.4%</b>
<b>April</b>	<b>4,951</b>	<b>9.2%</b>	<b>40,469</b>	<b>37.5%</b>	<b>22,892</b>	<b>18.9%</b>	<b>21,961</b>	<b>27.9%</b>	<b>23.4%</b>
<b>May</b>	<b>11,084</b>	<b>20.7%</b>	<b>10,438</b>	<b>9.7%</b>	<b>10,020</b>	<b>8.3%</b>	<b>10,464</b>	<b>13.3%</b>	<b>13.0%</b>
June	22,583	42.2%	47,247	43.8%	4,473	3.7%	20,811	26.4%	29.0%
July	11,608	21.7%	2,028	1.9%	13,417	11.1%	1,644	2.1%	9.2%
<b>June-July</b>	<b>34,191</b>	<b>63.8%</b>	<b>49,275</b>	<b>45.7%</b>	<b>17,890</b>	<b>14.8%</b>	<b>22,455</b>	<b>28.5%</b>	<b>38.2%</b>
"spring" (Mar-Jul)	50,309	93.9%	103,345	95.8%	58,194	48.0%	61,651	78.2%	
<b>TOTAL</b>	<b>53,574</b>	<b>100.0%</b>	<b>107,922</b>	<b>100.0%</b>	<b>121,142</b>	<b>100.0%</b>	<b>78,792</b>	<b>100.0%</b>	
<i>(hatchery smolts)</i>	26,130	32.8%	6,199	5.4%	4,660	3.7%	3,750	4.5%	

# Revised sampling proposal

## Five Proposed Strata

Five Proposed Strata							
1			2	3	4	5	
Dec			Mar	Apr	May	Jun	
Jan						Jul	
Feb							
ave. counts = 21,983			4,352	22,568	10,502	30,953	
ave. passage = unknown			17,755	129,061	80,447	476,875	
<i>sampling goals: 1,000</i>			<i>1,000</i>	<i>1,000</i>	<i>1,000</i>	<i>1,000</i>	

Baseline samples: 3 spring & 2 fall stocks

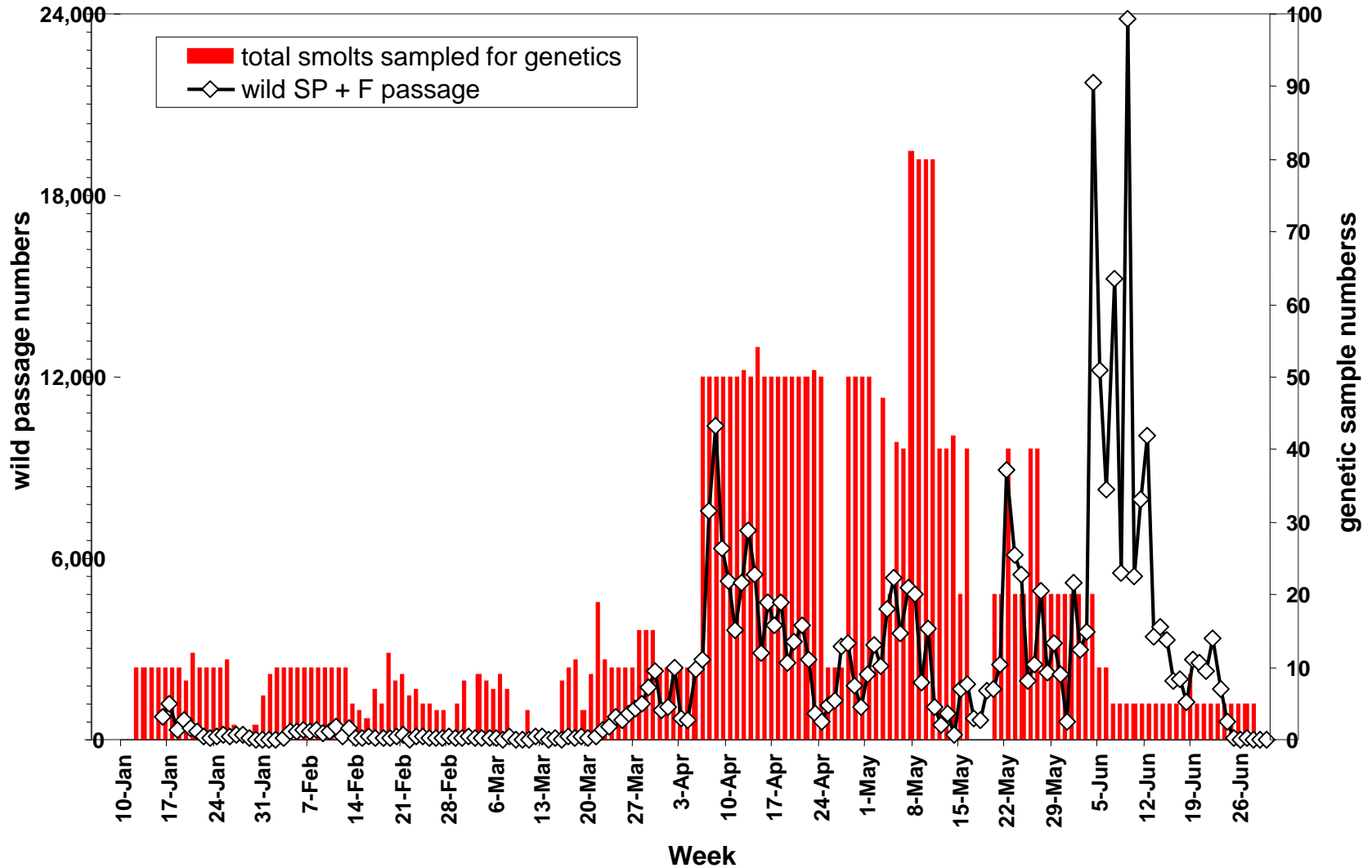
Unknown smolts - 1,500 (5 x 300)

Known-origin smolts: 150

# Chandler smolt sampling strategy

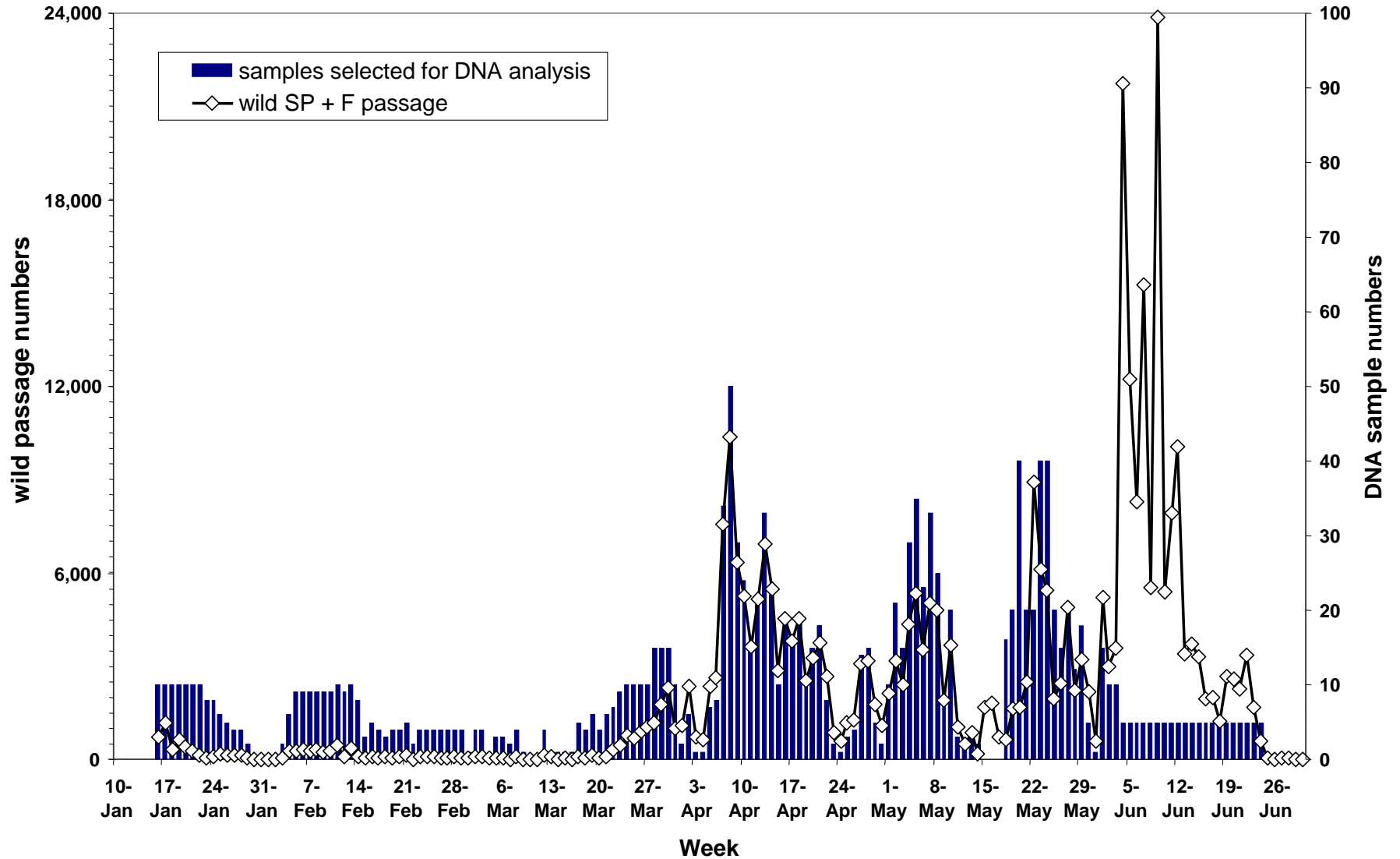
- Define daily trap abundance as either peak or non-peak
  - non-peak / peak thresholds variable between strata but constant within
  - high (Y) and low (X) abundance sampling goals vary between strata
  - $X < Y$
- If trap abundance estimate is below threshold, sample X number
- If trap abundance estimate is above threshold, sample Y number
- At end of season, plot daily sample numbers and estimated daily passage
- Sub-sample daily samples so that daily genetic sample proportions within strata approximate daily passage proportions within the same strata

# Jan – July smolt passage & genetic sampling



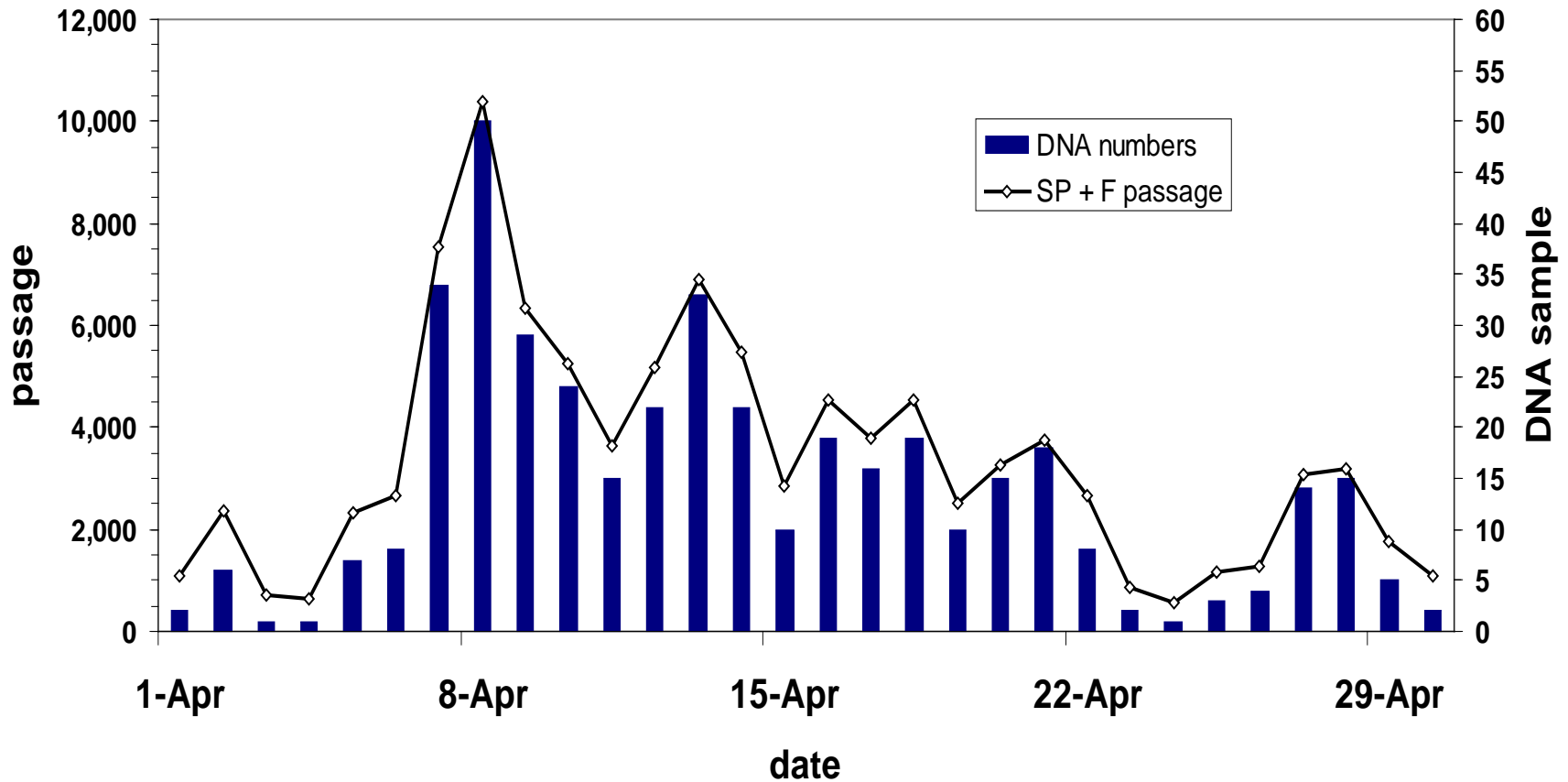


# Jan - July 2004 smolt passage & DNA sub-sampling



# April 2004 smolt passage & DNA sub-sampling

N = 415 analyzed; 1,186 sampled



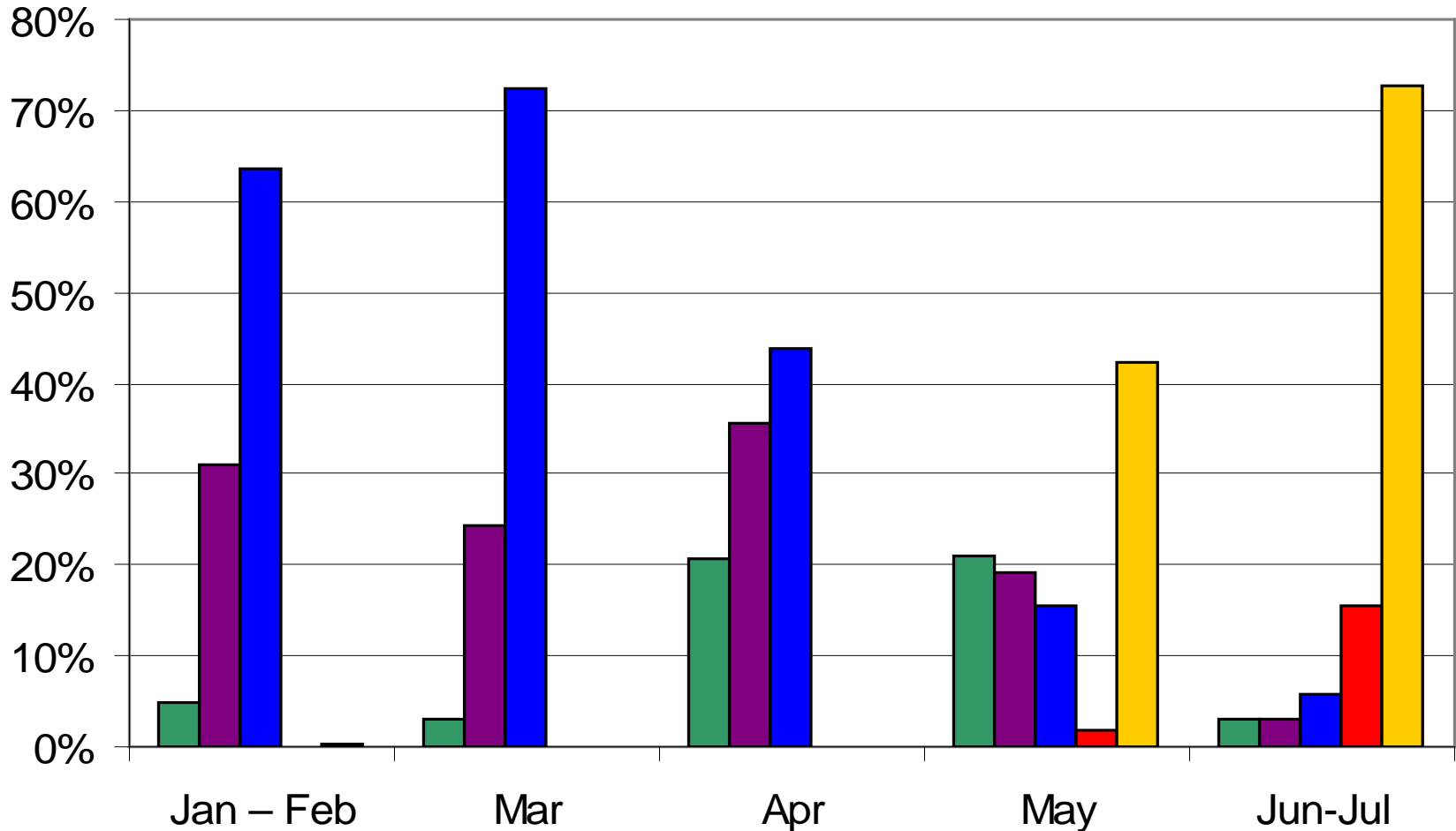
# Baseline collections for Chandler Chinook smolt analyses

Collection Source	Collection code	# processed
American River spring	89AG	80
	91DQ	102
	93DO	18
	<i>total</i>	<i>200</i>
Naches River spring	89AC	76
	89AI	26
	93DQ	50
	93DR	32
	<i>total</i>	<i>184</i>
upper Yakima River spring	92DN	24
	03GO	99
	<i>total</i>	<i>123</i>
Marion Drain fall	89BX	100
	92FQ	92
	93DY	8
	<i>total</i>	<i>200</i>
lower Yakima River fall	90DF	109
	93DW	82
	<i>total</i>	<i>191</i>

# Microsatellite loci used for 2004 Chandler Chinook smolt analysis

<b>Locus</b>	<b>Repeat (bp)</b>	<b>Allelic range (bp)</b>	<b>Observed # of alleles</b>
Oki-100	4	221 - 359	34
Ots-201b	4	169 - 338	41
Ots-208b	4	185 - 369	46
Ssa-408	4	211 - 332	33
Ogo-2	2	231 - 266	18
Ssa-197	4	170 - 318	38
Ogo-4	2	165 - 203	16
Ots-G474	4	183 - 247	16
Ots-213	4	209 - 385	43
Ots-3M	2	159 - 289	14
Ots-9	2	128 - 143	8
Ots-212	4	156 - 278	29

# Chinook smolt composition at Chandler - 2004



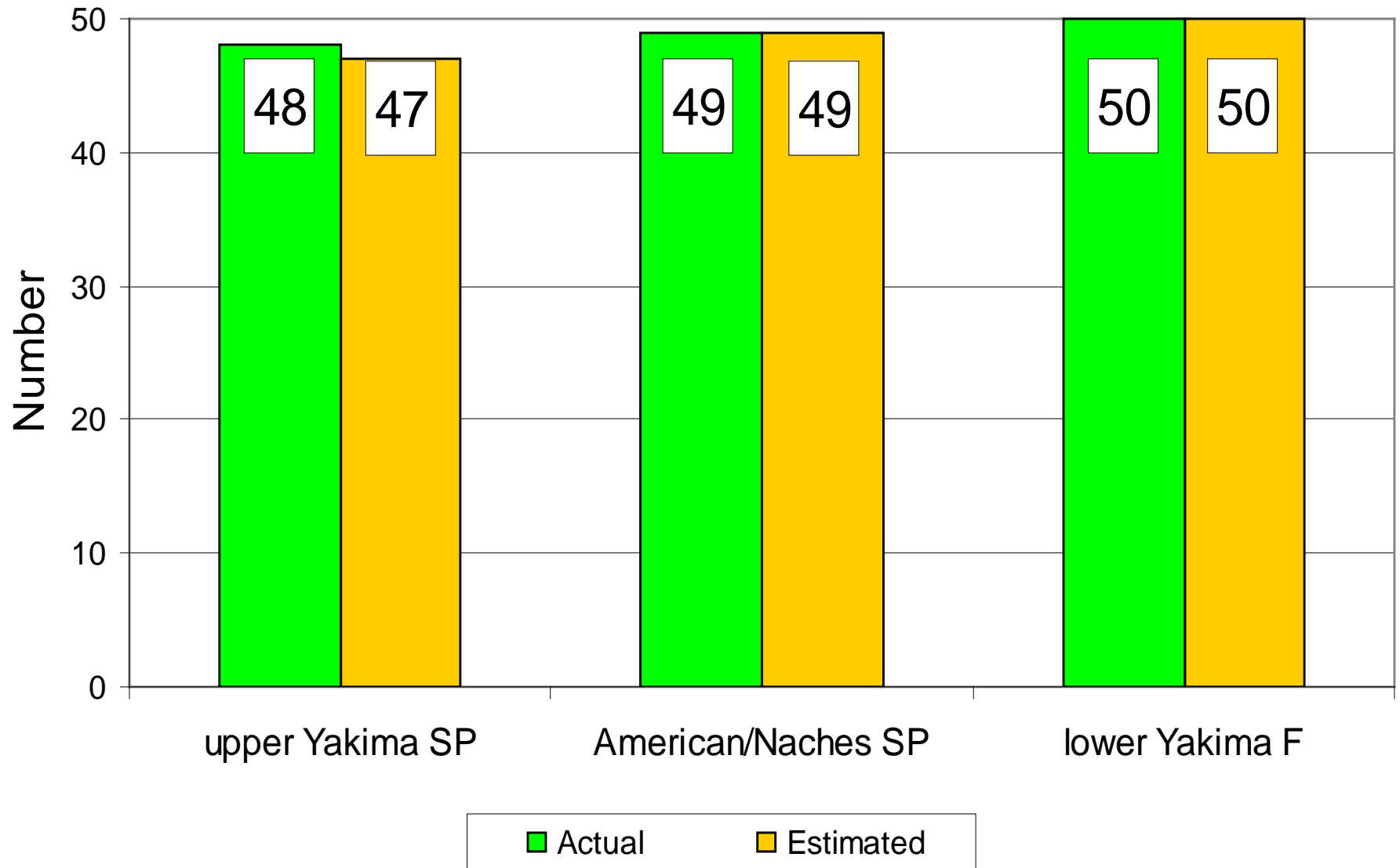
■ American R. SP ■ Naches R. SP ■ upper Yakima R. SP ■ Marion Drain F ■ lower Yakima R. F

*Estimated using MixAssign v1.1c - available from WDFW Genetics Lab*

# Blind test included to test method with real animals

- Yakama Nation samplers provided 150 known-origin samples
- WDFW Genetics Lab genotyped these as ***unknowns*** and assigned them to American/Naches spring, upper Yakima spring, or lower Yakima fall
- Genetics-based assignments were sent to Yakama Nation staff for comparison with points-of-capture

# Blind test estimated stock composition



# Blind test assignment accuracy

135/146 = 92% correct assignments

Actual \ Assigned	upper Yakima SP	American/Naches SP	lower Yakima F
upper Yakima SP	43	5	0
American/Naches SP	4	44	0
lower Yakima F	0	0	50



# Acknowledgements

**Funding provided by BPA**

**Yakama Nation samplers**

**Jennifer Von Bargen, Alice Frye, and Norm Switzler**



# Genetics-based mixture analyses

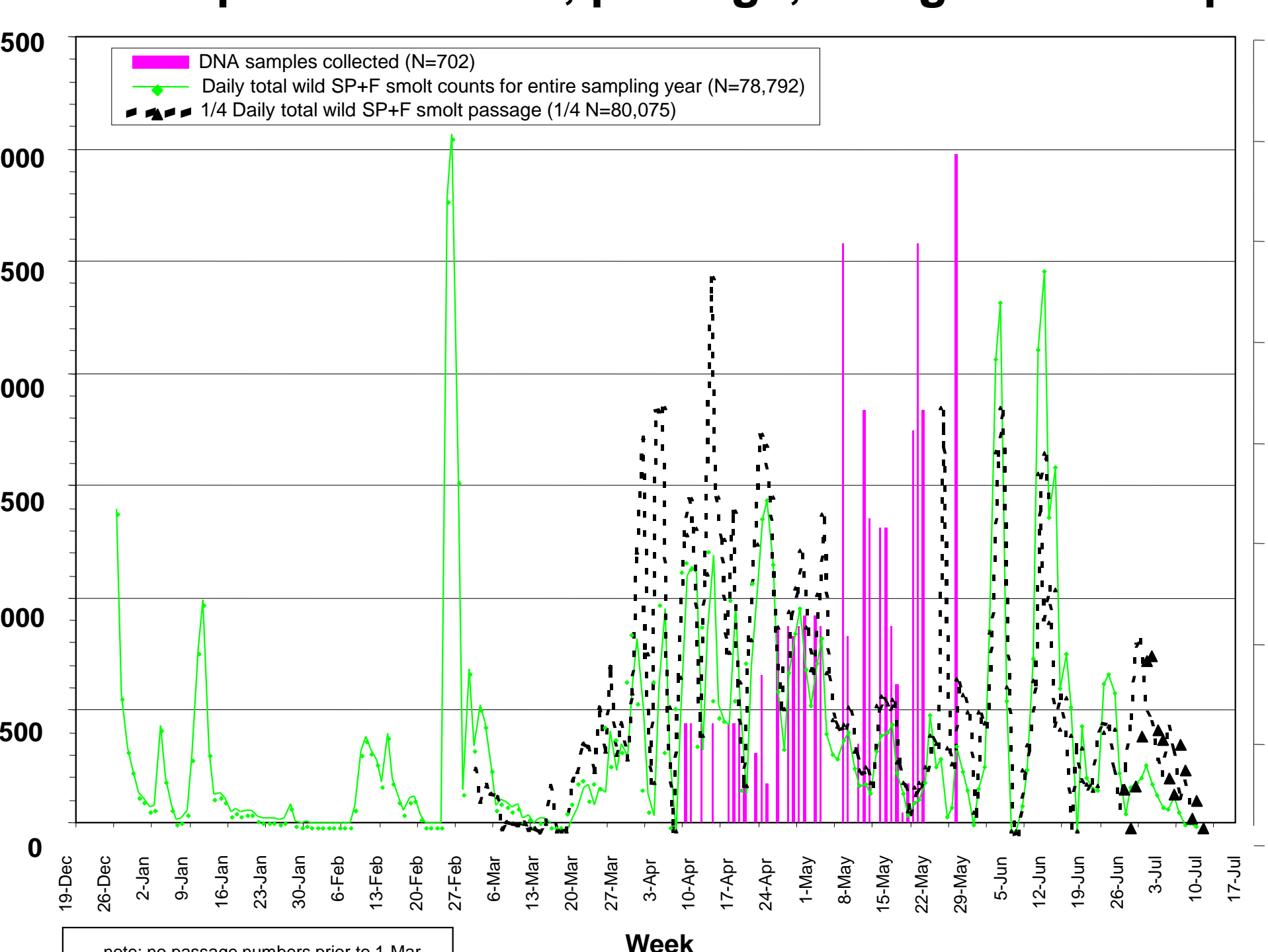
- Mixed stock analysis (MSA) – uses “baseline” data to estimate stock proportions in a mixture
  - “baseline” data are stock-specific allele frequencies
  - MSA estimates the likely stock composition of the mixture sample drawn from a sampling stratum

# Genetics-based mixture analyses

- Mixed stock analysis (MSA) – uses “baseline” data to estimate stock proportions in a mixture
- Individual assignment – uses “baseline” data to estimate stock proportions and assign individuals to stock-of-origin simultaneously
  - “baseline” data are stock-specific allele frequencies
  - estimates the likely stock composition of the mixture sample drawn from a sampling stratum
  - identifies the most likely stock-of-origin of each individual within a mixture sample

# Genetics-based mixture analyses

- Mixed stock analysis (MSA) – uses “baseline” data to estimate stock proportions in a mixture
- Individual assignment – uses “baseline” data to estimate stock proportions and assign individuals to stock-of-origin simultaneously
- Clustering methods – attempt to identify genotypic clusters (= stocks?) in a mixture without assuming knowledge of stock structure



# Results of 1000 bootstrapped simulations of estimated stock proportions for the 4 strata

<b>Estimated % - January February strata</b>						
	N =	u Yakima R. SP	Naches R. SP	American R. SP	Marion Drain F	l Yakima R. F
u Yakima R. SP	<b>168</b>	<b>164.37</b>	<b>3.45</b>	<b>0.15</b>	<b>0.01</b>	<b>0.02</b>
Naches R. SP	<b>82</b>	<b>37.11</b>	<b>43.43</b>	<b>1.46</b>	<b>0</b>	<b>0</b>
American R. SP	<b>13</b>	<b>5.98</b>	<b>0.73</b>	<b>6.29</b>	<b>0</b>	<b>0</b>
Marion Drain F	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
l Yakima R. F	<b>1</b>	<b>0.31</b>	<b>0</b>	<b>0</b>	<b>0.09</b>	<b>0.6</b>
<b>Estimated % - April strata</b>						
	N =	u Yakima R. SP	Naches R. SP	American R. SP	Marion Drain F	l Yakima R. F
u Yakima R. SP	<b>182</b>	<b>176.02</b>	<b>5.44</b>	<b>0.52</b>	<b>0.01</b>	<b>0.02</b>
Naches R. SP	<b>147</b>	<b>62.3</b>	<b>78.87</b>	<b>5.82</b>	<b>0.01</b>	<b>0</b>
American R. SP	<b>86</b>	<b>38.01</b>	<b>1.96</b>	<b>46.04</b>	<b>0</b>	<b>0</b>
Marion Drain F	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
l Yakima R. F	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>

<b>Estimated % - March strata</b>						
	N =	u Yakima R. SP	Naches R. SP	American R. SP	Marion Drain F	l Yakima R. F
u Yakima R. SP	<b>120</b>	<b>118.12</b>	<b>1.78</b>	<b>0.08</b>	<b>0.01</b>	<b>0.01</b>
Naches R. SP	<b>41</b>	<b>19.43</b>	<b>20.89</b>	<b>0.69</b>	<b>0</b>	<b>0</b>
American R. SP	<b>5</b>	<b>2.26</b>	<b>0.32</b>	<b>2.42</b>	<b>0</b>	<b>0</b>
Marion Drain F	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
l Yakima R. F	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>Estimated % - June July strata</b>						
	N =	u Yakima R. SP	Naches R. SP	American R. SP	Marion Drain F	l Yakima R. F
u Yakima R. SP	<b>8</b>	<b>7.63</b>	<b>0.26</b>	<b>0.03</b>	<b>0</b>	<b>0.09</b>
Naches R. SP	<b>4</b>	<b>0.37</b>	<b>2.01</b>	<b>0.24</b>	<b>0</b>	<b>1.38</b>
American R. SP	<b>4</b>	<b>0.02</b>	<b>0.08</b>	<b>2.18</b>	<b>0</b>	<b>1.73</b>
Marion Drain F	<b>22</b>	<b>0.03</b>	<b>0</b>	<b>0</b>	<b>5.94</b>	<b>16.03</b>
l Yakima R. F	<b>102</b>	<b>0.02</b>	<b>0</b>	<b>0</b>	<b>1.33</b>	<b>100.65</b>

# Jackknifed baseline samples

From \ To	upper Yakima R. SP	Naches R. SP	American R. SP	Marion Drain F	lower Yakima F
upper Yakima R. SP	84.6%	13.0%	0.0%	1.6%	0.8%
Naches R. SP	16.3%	64.7%	15.2%	0.5%	0.0%
American R. SP	2.0%	7.0%	89.0%	0.0%	0.0%
Marion Drain F	2.5%	0.5%	0.0%	65.0%	31.0%
lower Yakima F	0.5%	0.5%	0.0%	26.7%	72.3%



# Terminology

- locus – the genetic address of a “gene”
- allele – a form of the “gene” at a locus
- genotype – the allelic composition of an individual
- allele frequency – the relative abundance of an allele in a group or population
- microsatellite – a class of locus typified by repeating DNA sequence