



Advantages and Limitations of Genetic Stock Identification Applied to Pink Salmon Stocks


Andres Araujo, John Candy, and Terry Beacham
Molecular Genetics Laboratory
Fisheries and Oceans Canada

25th Pink and Chum
Salmon Workshop

2012

Juneau, AK

Genetic Stock Identification (GSI)

- What is GSI?
- Advantages of GSI
- Software (**Oncor** ,  Cbayes.Ink)
- GSI in management
- Limitations of GSI

Pinks more difficult to identify ?

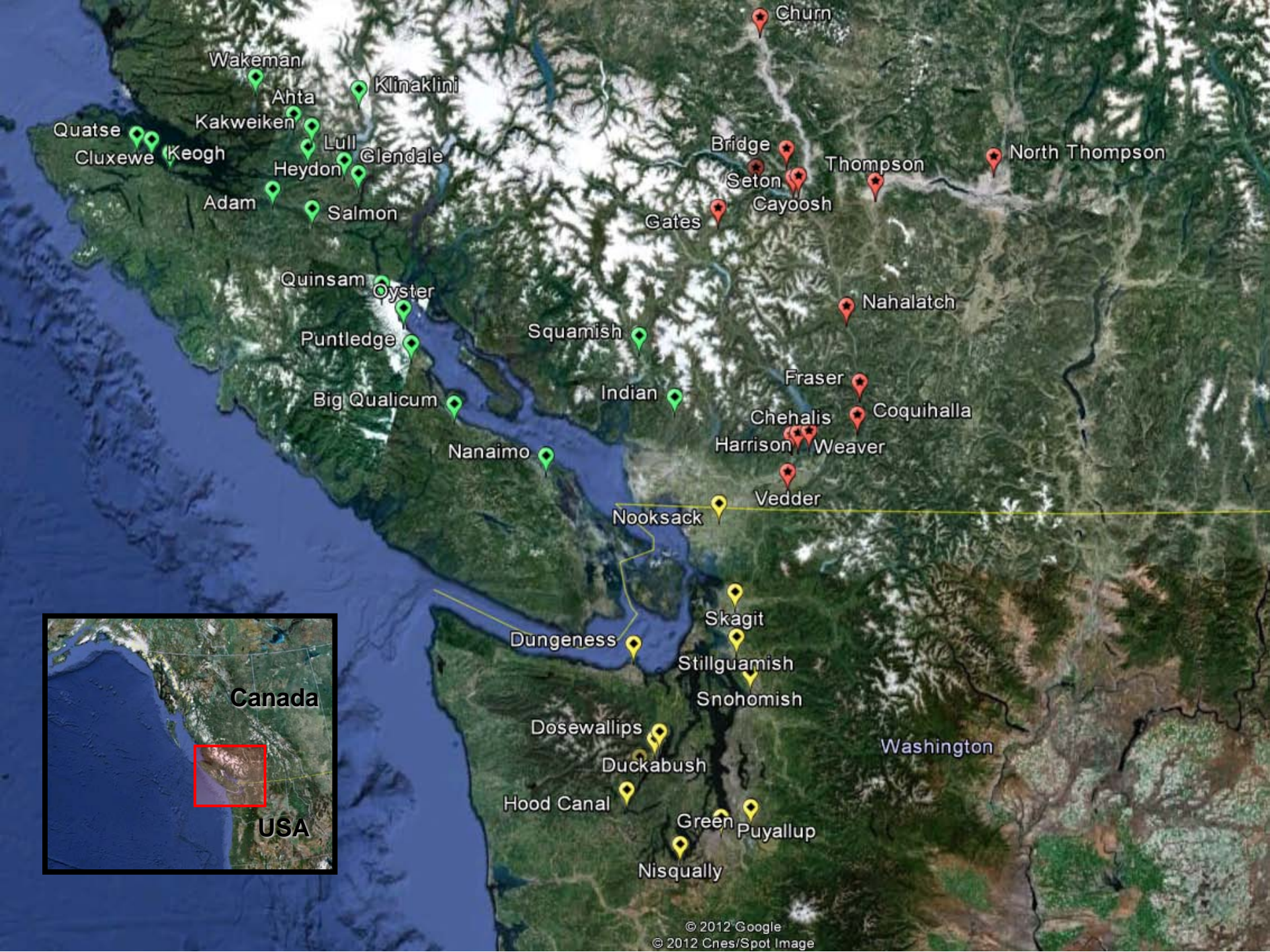


Pinks more difficult to identify ?

- Faithfulness to spawning site ?
 - Some populations stray; others don't
- Overall interest in the species (sampling)?
 - Smallest DFO baseline of 5 major salmonids
- Genetic separation among population (F_{st})
 - Variable no. of selected (16 vs.13) loci increased regional accuracy (Beacham et al *in press*)

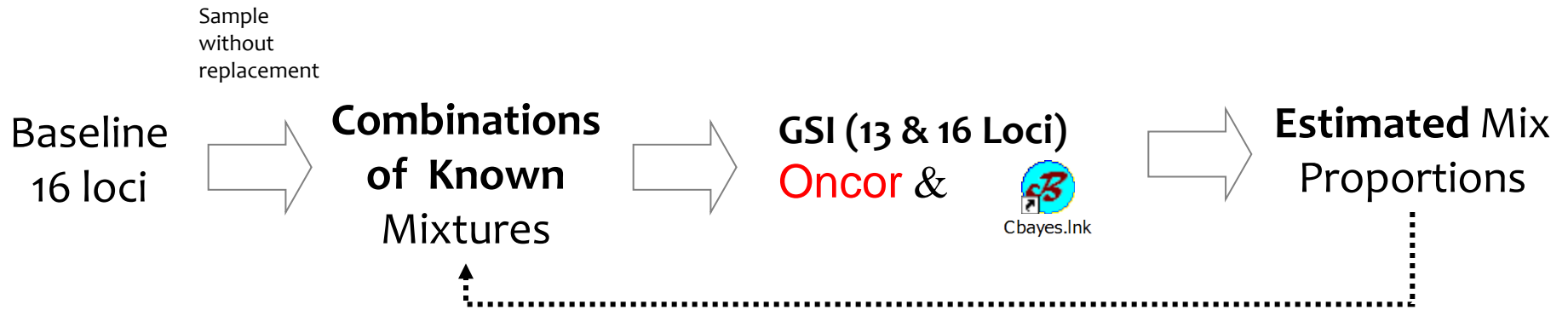
Objectives

1. Test if using a variable number of selected loci increases within-region and stock specific accuracy (real data).
2. Explore the limitations of GSI, identify gaps and biases of software, and propose corrective mechanisms (simulated data).



Methods Objective 1

Testing variable number of selected loci



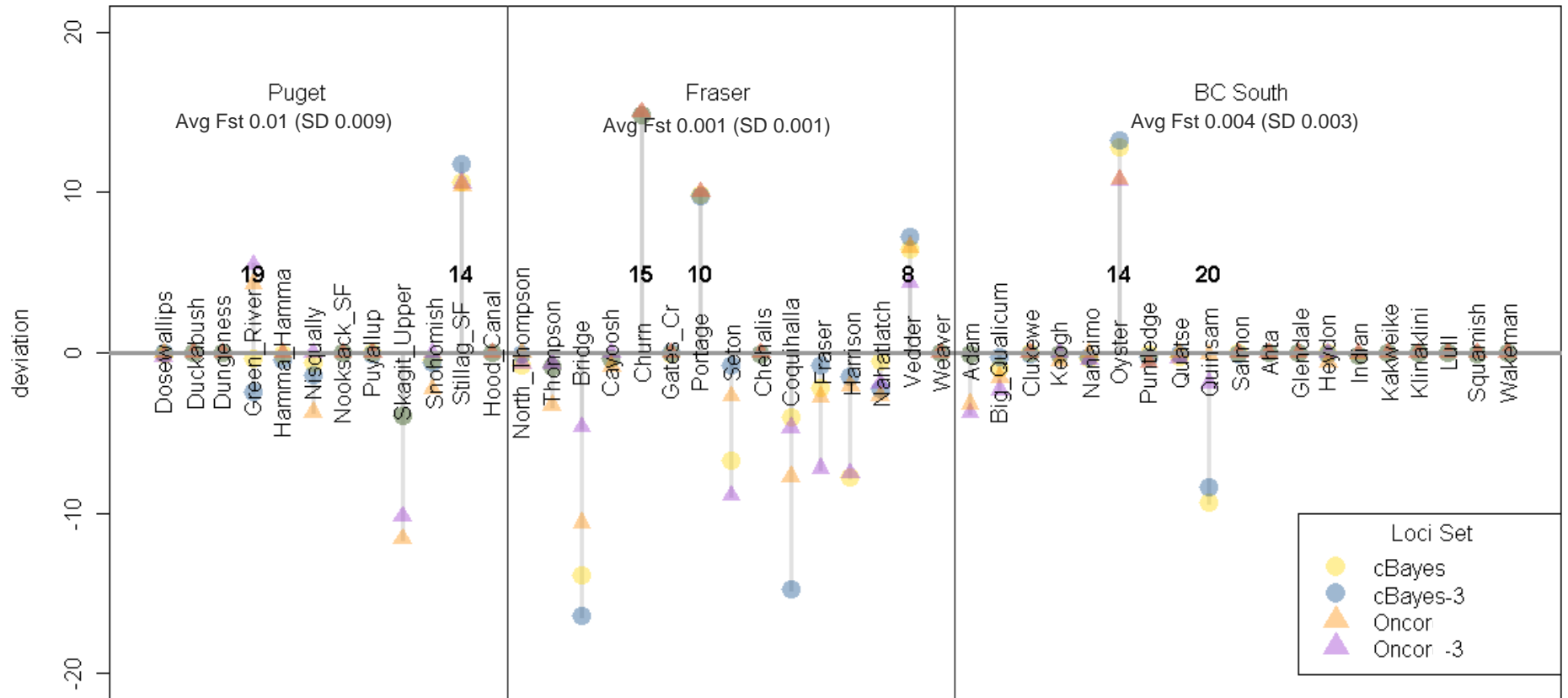
Kalinowski, S. T., K. R. Manlove, and M. L. Taper. 2007. ONCOR A computer program for Genetic Stock Identification. . Department of Ecology, Montana State University, Bozeman MT 59717. Available for download from <http://www.montana.edu/kalinowski>

Neaves, P. I., C. G. Wallace, J. R. Candy, and T. D. Beacham. 2005. CBayes: Computer program for mixed stock analysis of allelic data. Version v5.01.

Results Objective 1

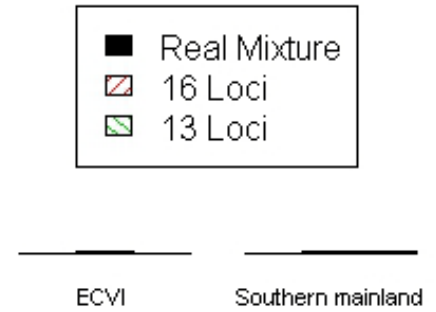
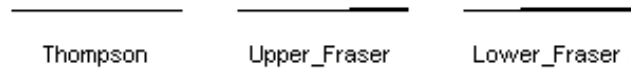
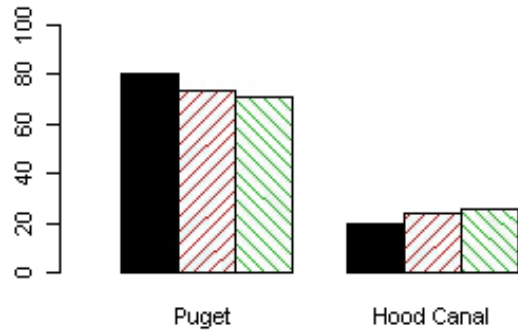
Testing variable number of selected loci

PLOT OF DEVIATIONS ~1/3 EACH REGION

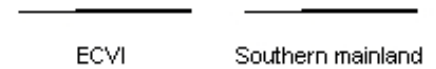
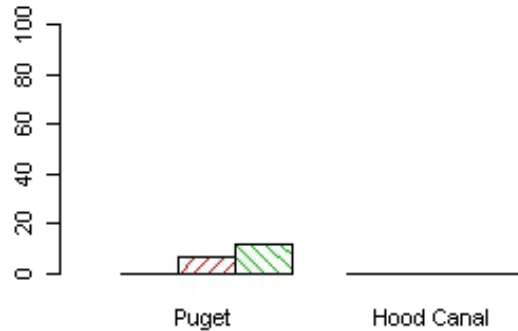


Within-Region Mixtures

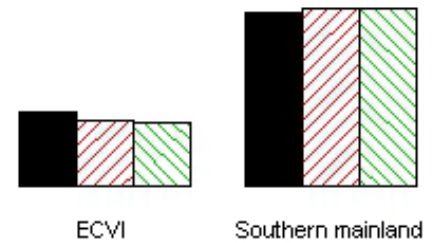
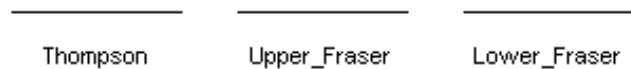
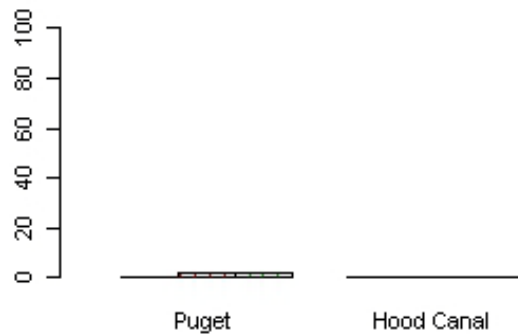
Hood Ch. 20% - Other Puget 80%



Upper 50% - Lower 50% Fraser

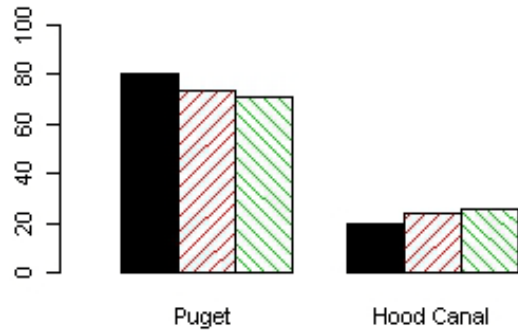


ECVI 30% - Mainland 70%

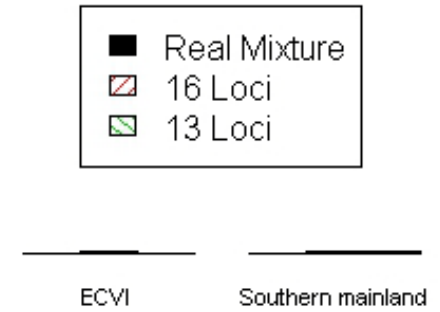


Within-Region Mixtures

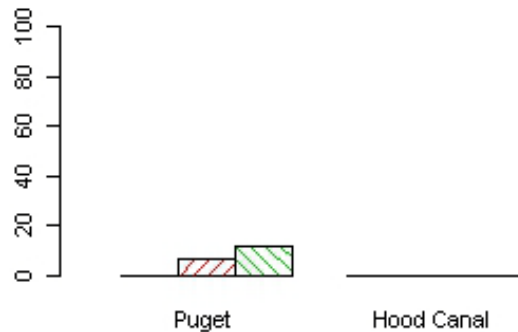
Hood Ch. 20% - Other Puget 80%



Thompson Upper_Fraser Lower_Fraser

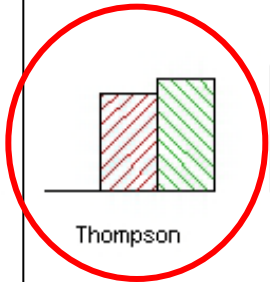


Upper 50% - Lower 50% Fraser

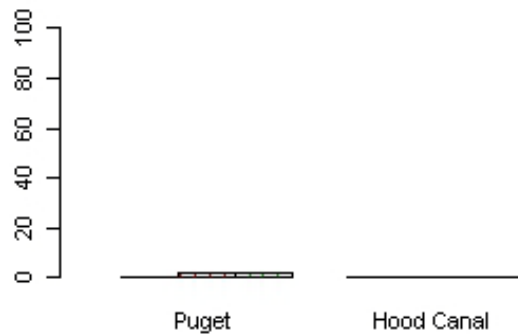


Thompson Upper_Fraser Lower_Fraser

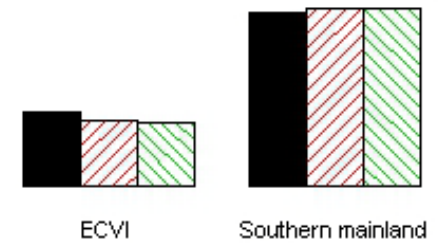
ECVI Southern mainland



ECVI 30% - Mainland 70%

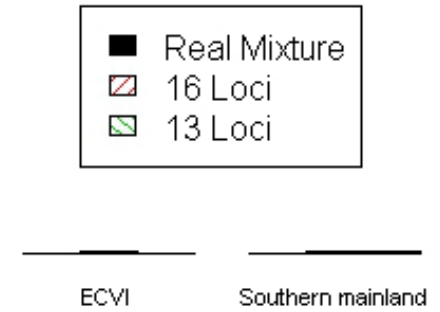
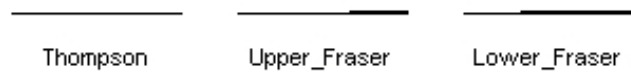
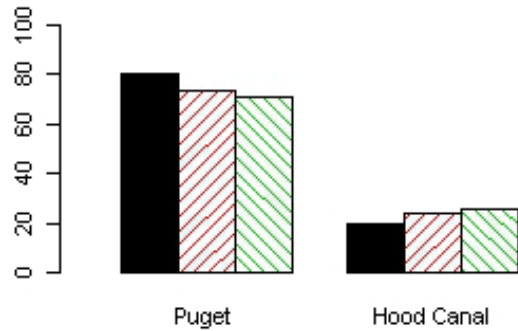


Thompson Upper_Fraser Lower_Fraser

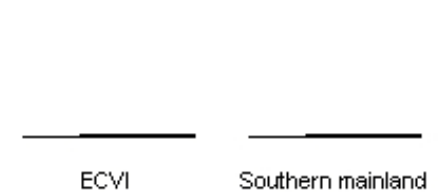
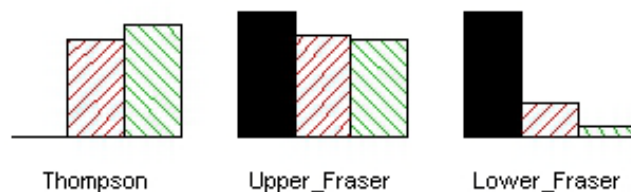
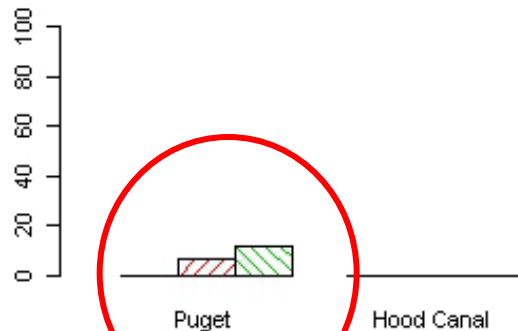


Within-Region Mixtures

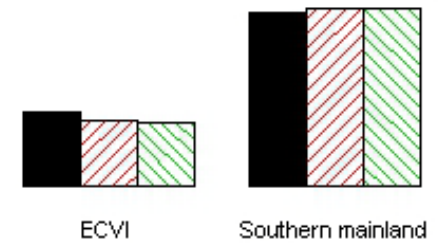
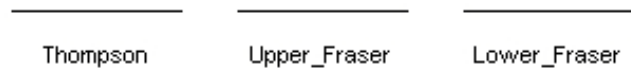
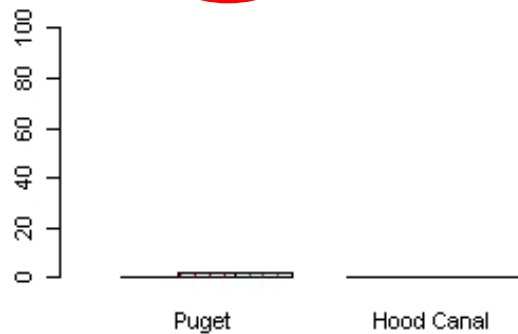
Hood Ch. 20% - Other Puget 80%



Upper 50% - Lower 50% Fraser



ECVI 30% - Mainland 70%

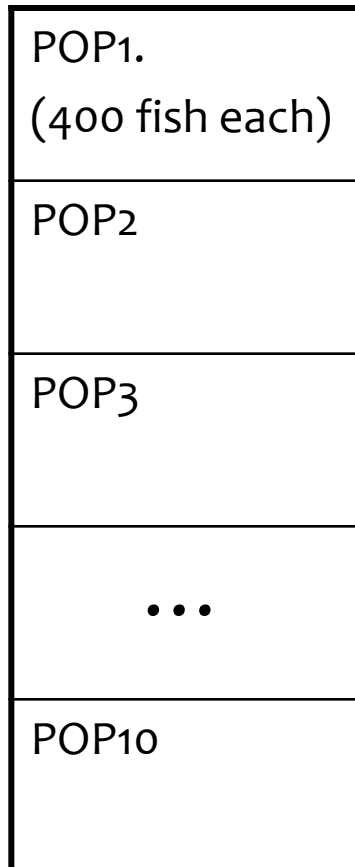


Methods Objective 2

Exploring GSI Limitations

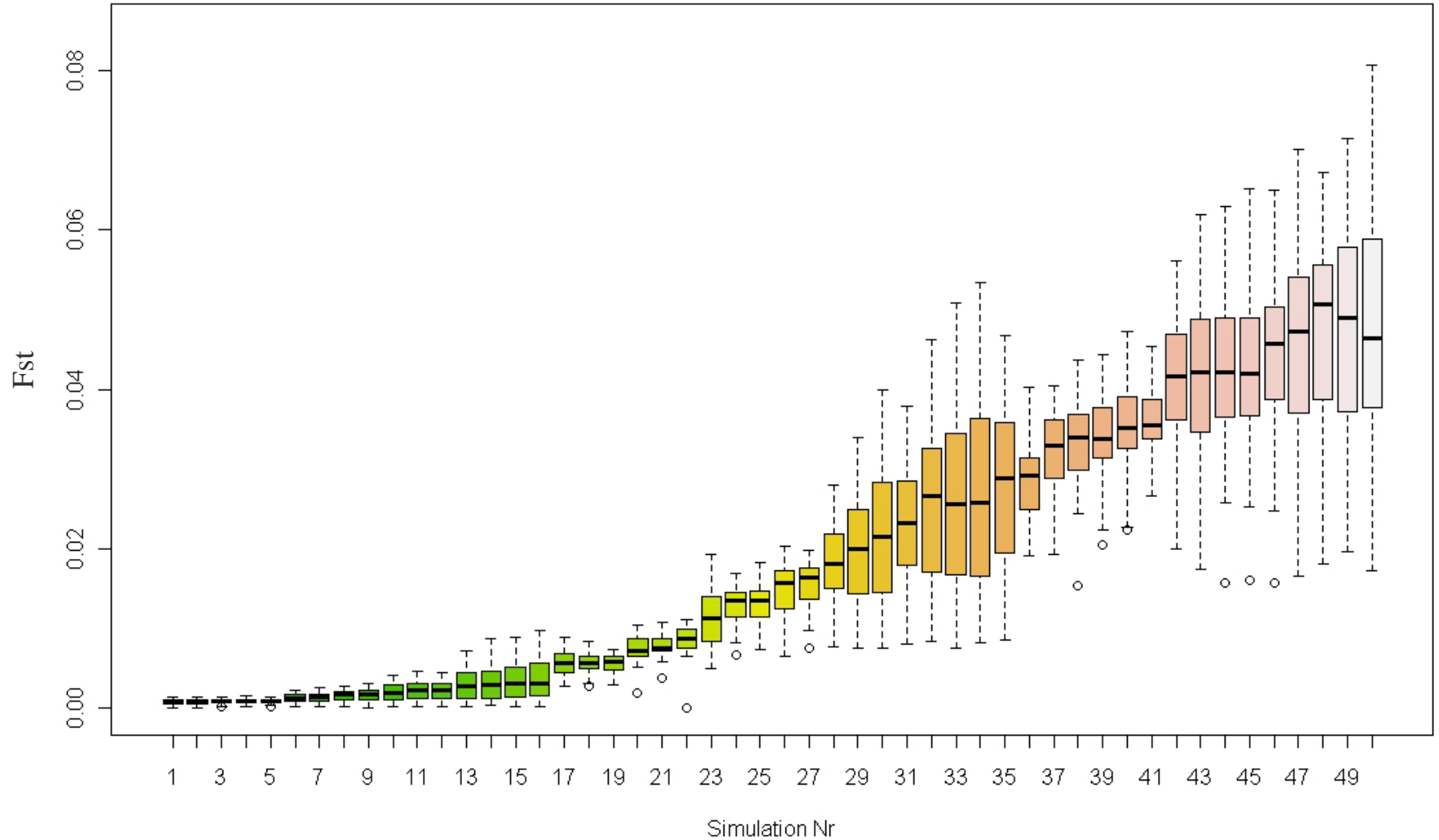
50 baselines *EASYPop*
(16 loci, Avg 48 alleles)

0.0007 to 0.04 AVG *F_{st}*s



Simulated Populations

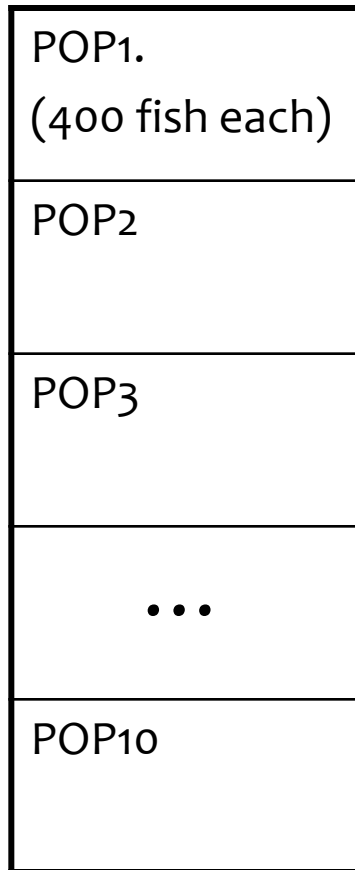
Simulated Baselines in EasyPOP



Methods Objective 2

Exploring GSI Limitations

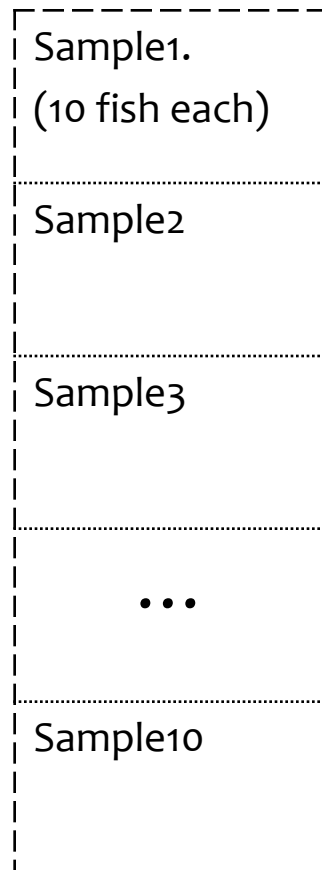
50 baselines *EASYPop*
(16 loci, Avg 48 alleles)
0.0007 to 0.04 AVG *F_{st}*s



Sample without replacement



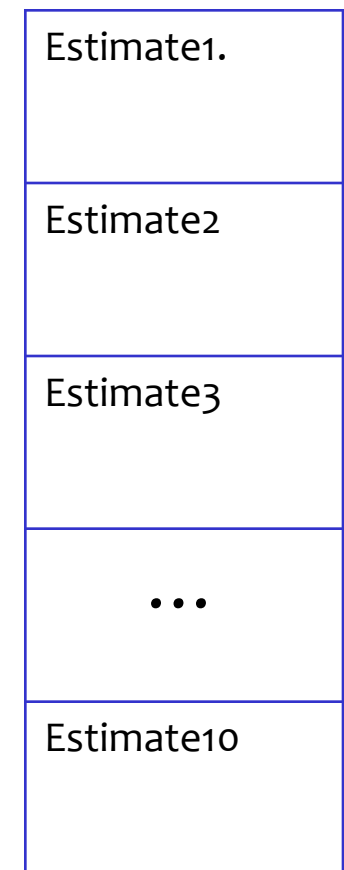
Known
Mixtures



Oncor



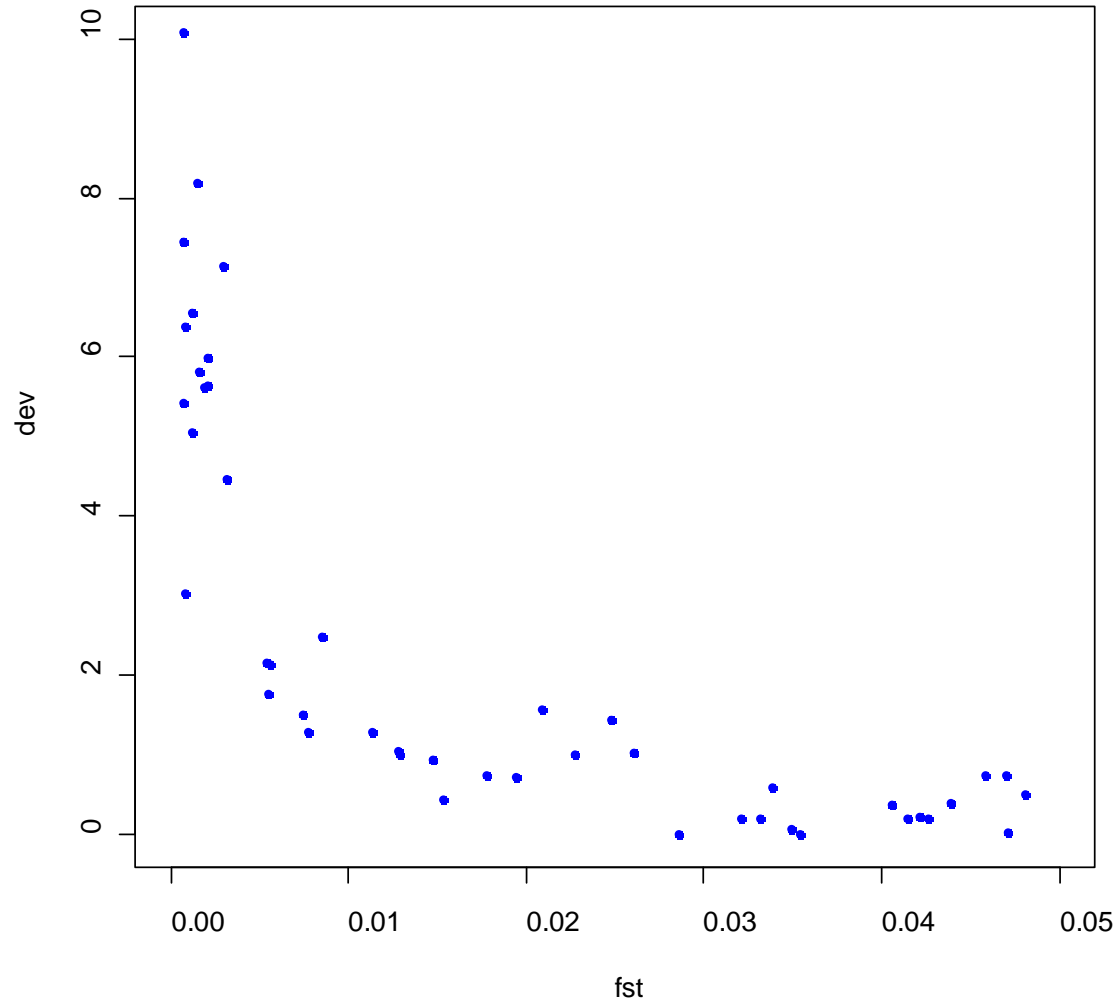
Estimated Mix
proportions



Results

Exploring GSI limitations

CBayes vs Oncore and GSI accuracy



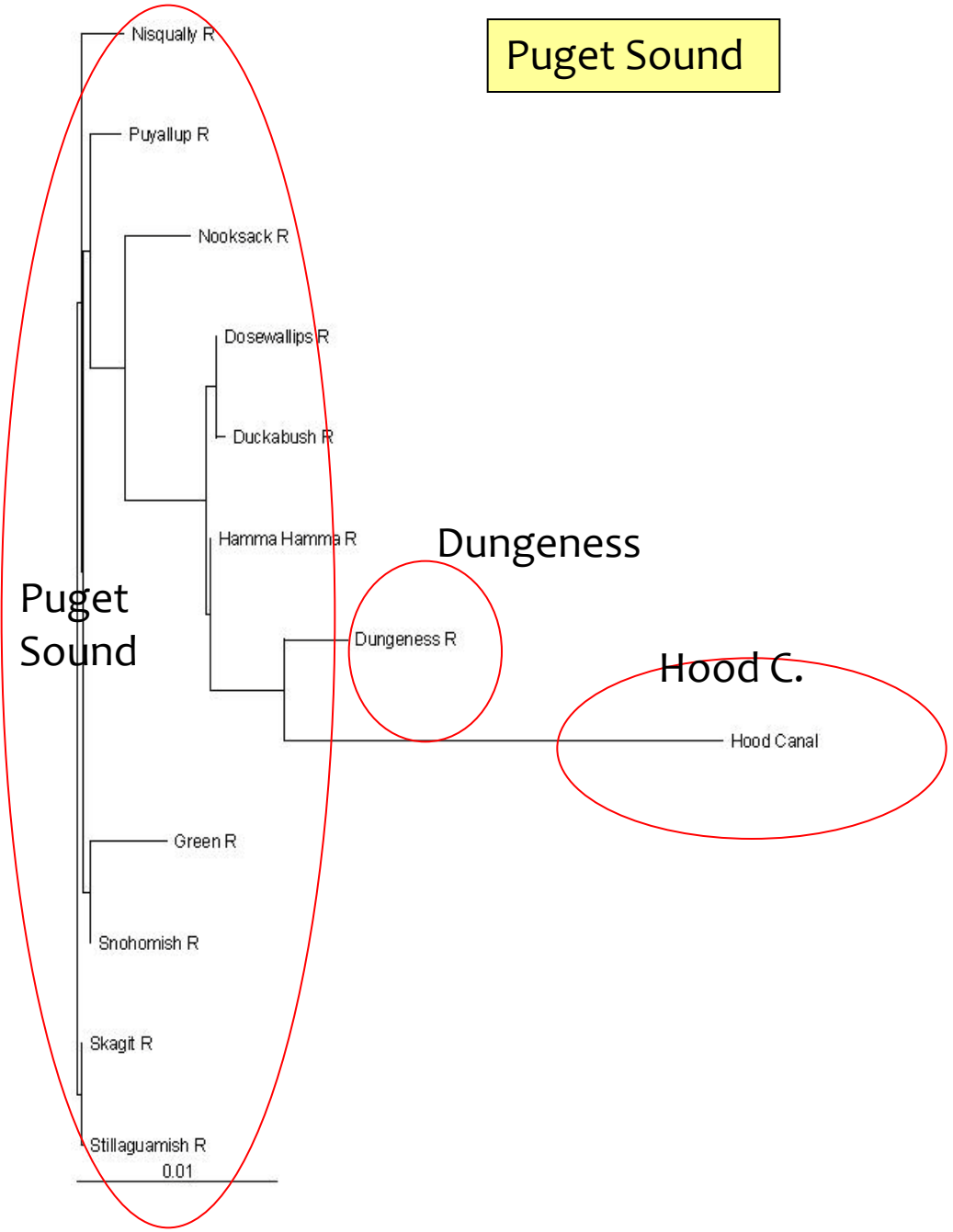
Conclusions

- Using more or fewer Loci does not increase the stock-specific or within-region definition. Using 16 Loci increases the among-region accuracy, current management application.
- For GSI purposes, very distant populations may belong to the same stock. Neighbouring populations may be genetically very different.
- GSI subject to large errors and biases when F_{st} s are small. Simulation shows GSI very inaccurate (more than 20% error) when $F_{st} < 0.01$.

Conclusions

- GSI is limited in two ways: mathematically (failure to identify stocks at very low F_{st} s), and biologically (assumes that populations are highly faithful to spawning sites).
- Software ONCOR performs better at lower F_{st} s (< 0.01) than CBAYES. No difference in performance at larger F_{st} s.
- It is crucial to assess the F_{st} in Baseline populations and combine genetically closely related stocks when possible ($Avg F_{st} > 0.01$).

Puget Sound



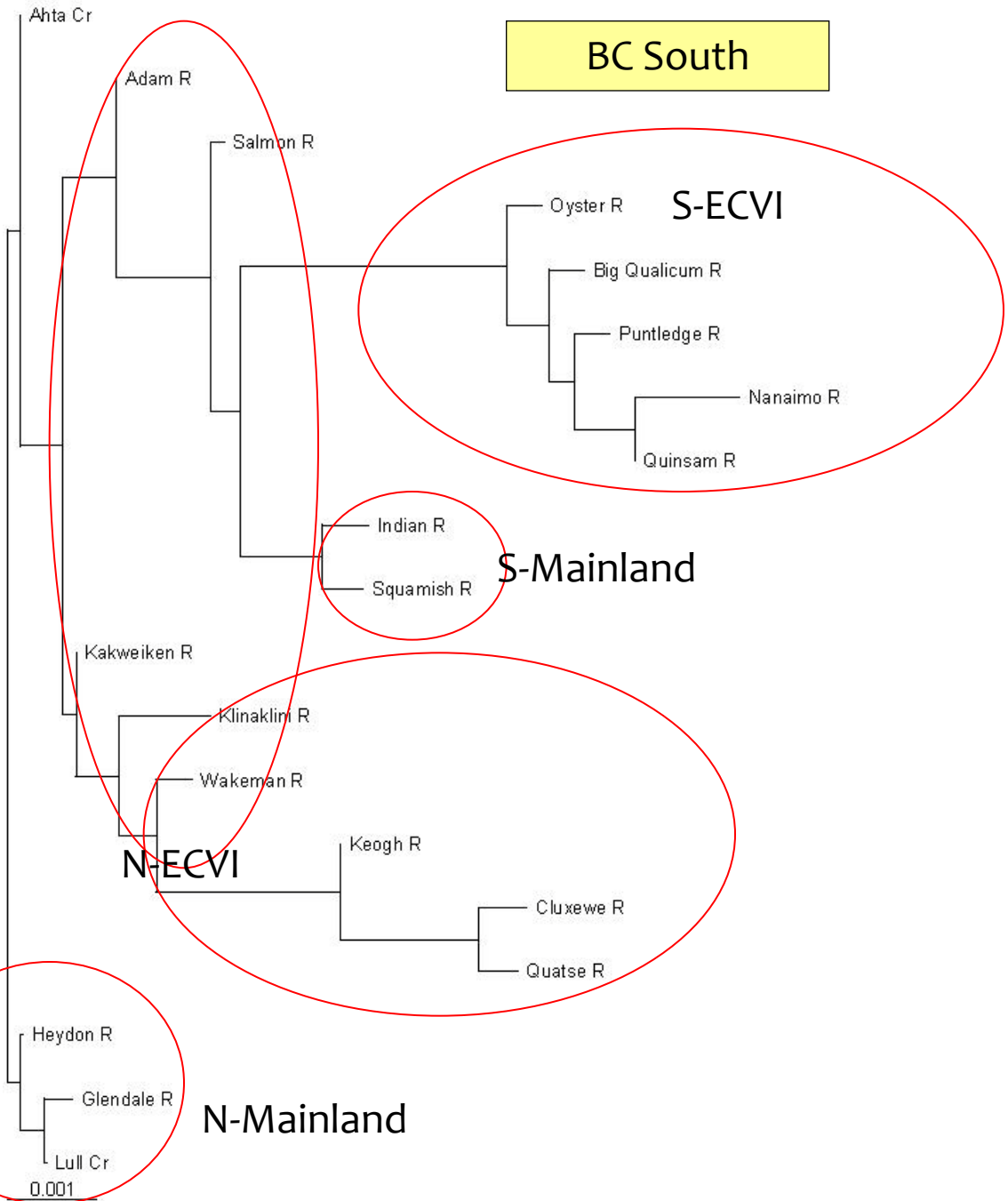
Puget
Sound

Dungeness

Hood C.

0.01

BC South



S-ECVI

S-Mainland

N-ECVI

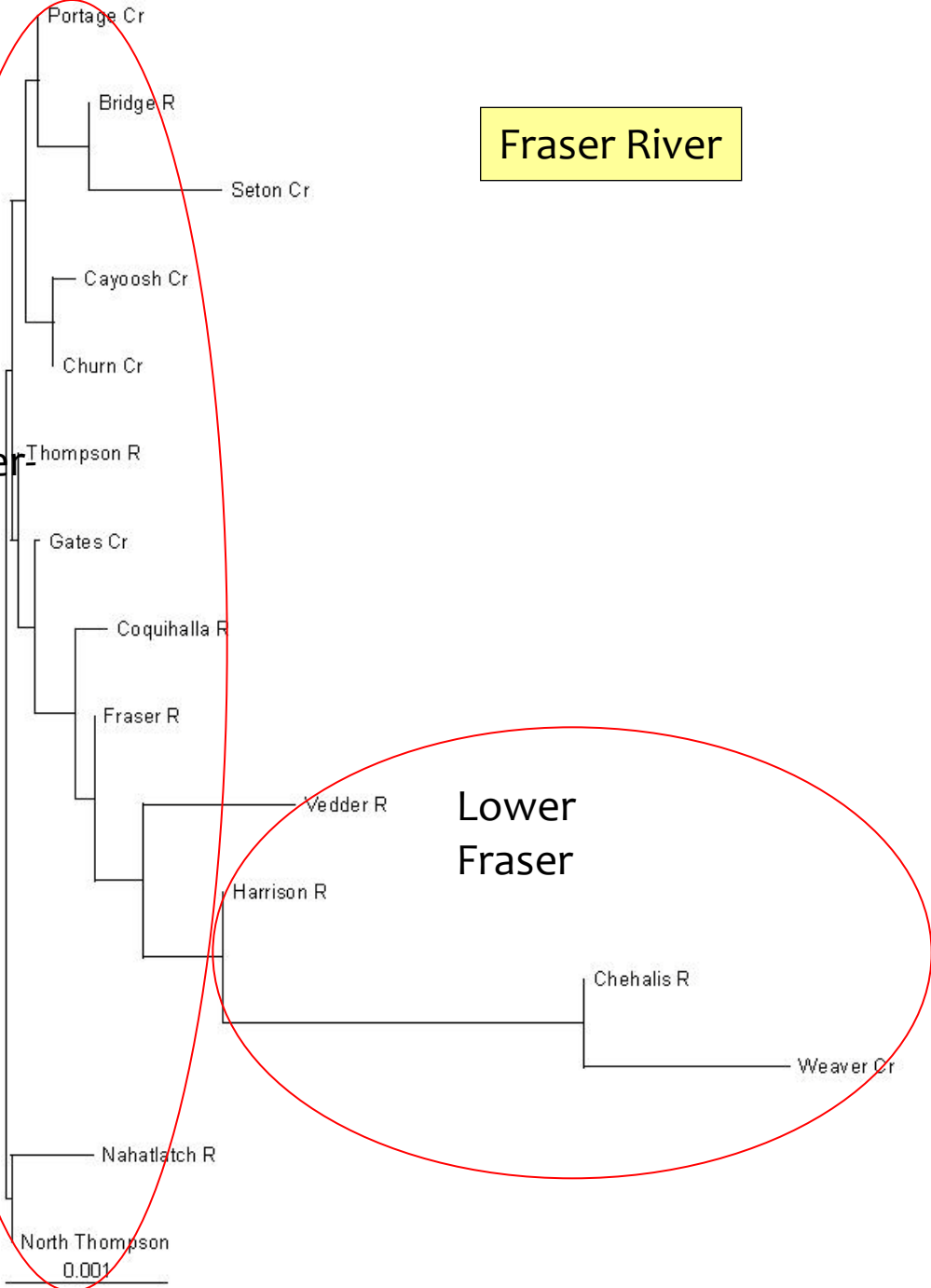
N-Mainland

0.001

Fraser River

Upper Fraser
Thompson

Lower
Fraser



0.001

Acknowledgements



Lab Staff at MGL

Carrie Gummer
Katherine Horst
Kim Jonsen
Khai D. Lee
Cathy MacConnachie
Brenda McIntosh
Liane Stenhouse
Amy Tabata
Debra Tuck
Michael Wetklo
Colin G. Wallace

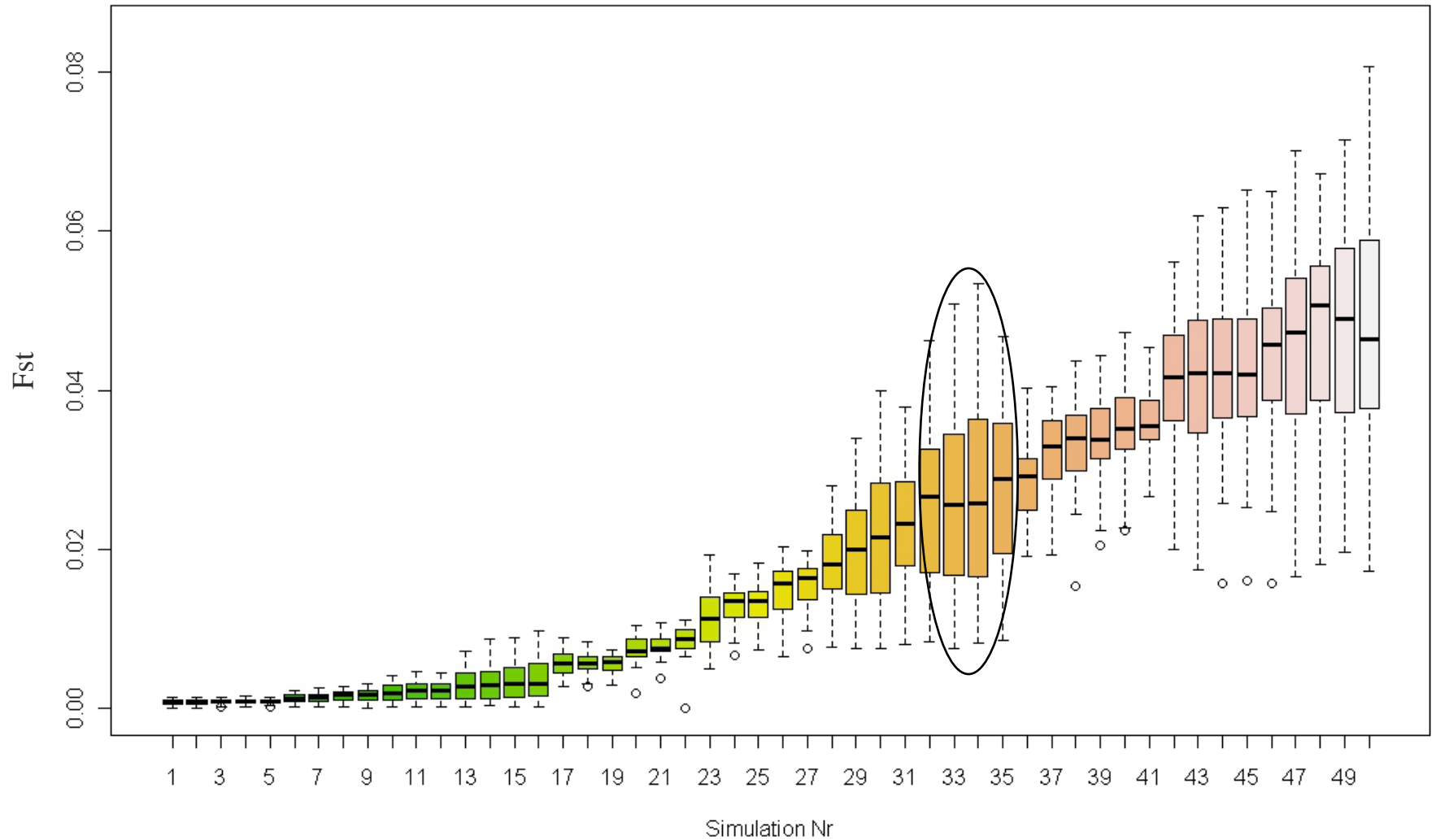


Pacific Salmon Commission

Steve Latham
Bruce White

Simulated Populations

Simulated Baselines in EasyPOP



Easy POP

Mating System	
proportion of female migration within groups?(between 0 and 1)	
proportion of male migration within groups?(between 0 and 1)	
proportion of female migration between groups?(between 0 and 1)	
proportion of male migration between group?(between 0 and 1)	
Mutation rate	
Number of generations?	
Ploidy level? (0=haplo-diploid; 1=haploid; 2=diploid)	diploid
Two sexes?:y/n	yes
Number of populations?	10
Same number of individuals in each population ?:y/n	yes
Number of females in each population?	200
Number of males in each population?	200
Same migration scheme over all simulation? (y/n)	yes
Migration model?	hierarchical stepping stone
Number of loci?	16
Free recombination between loci?: y/n	yes
Do all loci have the same mutation scheme?: y/n	yes
Mutation model	Ssm
Number of possible allelic states?(below 1000)	48
Variability of the initial population	max