Parentage-based tagging as an alternative to CWTs for salmon fishery management

John Carlos Garza & Eric C. Anderson
Southwest Fisheries Science Center
Santa Cruz CA

Current CWT tagging system

- Mechanically inserted and manually extracted metal tags that are manually read under a microscope
- Since 1968, 71 agencies in 5 states and B.C. have inserted ~600 miles of wire and tagged ~ 1 billion salmon and steelhead
- Until 1996, only fish with CWTs generally received adipose fin clips
- Nearly 1 million heads analysed at Juneau head lab alone.

Images from www.nmt.us

Current CWT tagging system

- Very useful tagging system over its 30+ year life
- Crucial to PSC objective of estimating fishery mortality on multiple stocks
- Provides stock of origin AND **cohort** of origin
- Large historical databases of tag recoveries provide comparative baseline

Challenges to CWT system

- Very low tag recovery rates (1.6 per 1,000 in chinook)
- Tag loss rates are poorly known
- CWT harvest may be underreported
- Mass-marking - Not all Ad-clipped fish have CWTs
- Assumption of equivalency of hatchery indicator stocks and genetically similar naturally spawning stocks (the gorilla assumption) - can be large areas.

Parentage-based Tagging

(a.k.a. the method formerly known as FPG)

- Highly efficient, transgenerational, genetic tagging method
  - Genotype all hatchery parents
  - Create reference (parent) database of all possible parent pairs
  - Fishery sampling and genotyping in offspring generation
  - Query of reference (parent) database to determine if parents are present
  - Determine parental pair and, therefore, hatchery stock of origin and exact age
- Information obtained for each tag recovery is the same as for a CWT (+more)
- By genotyping two parents, you effectively tag all of their 1,000s of offspring
- Requires no juvenile tagging, but MUCH higher tagging rates feasible.

Parentage-based Tagging

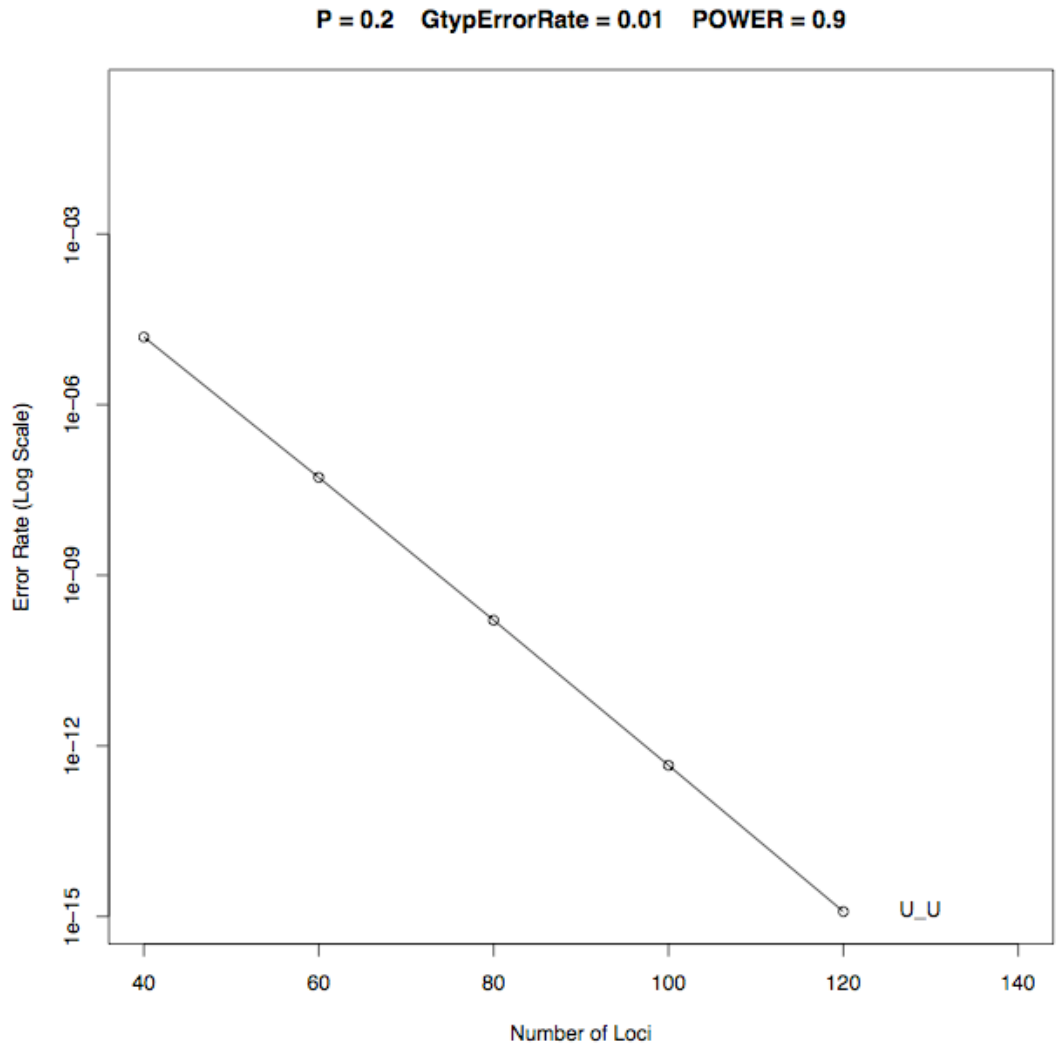
(a.k.a. the method formerly known as FPG)

- Fundamentally different than genetic stock identification: matches fishery sample to pairs of parents in reference (parent) database that have Mendelian compatiblity. GSI uses frequency based probability assessments
- Can be done using either traditional exclusion or maximum likelihood
- Power comes from number of loci, since each locus is an opportunity for incompatibility
- Marking and sampling issues with other tagging systems don't entirely go away.

Power analysis for large scale parentage-inference

- Anderson and Garza (2006; Genetics) evaluated the plausibility of large scale PBT through evaluation of power of SNP markers to infer parentage
  - Determine false positive rates in large scale parentage inference studies
  - Evaluate number of SNP loci necessary to correctly ID parent pairs
  - Describe new analytical method for fast ML parentage analysis
  - Evaluate effects of allele frequency, genotyping error and presence of kin
- 100 SNP genotype can identify parental pairs with false positive rate less than 1 fish per 300,000 fishery samples.
- False positive rates decrease exponentially with number of loci!

Exponential decline in false positive rate with number of SNP loci

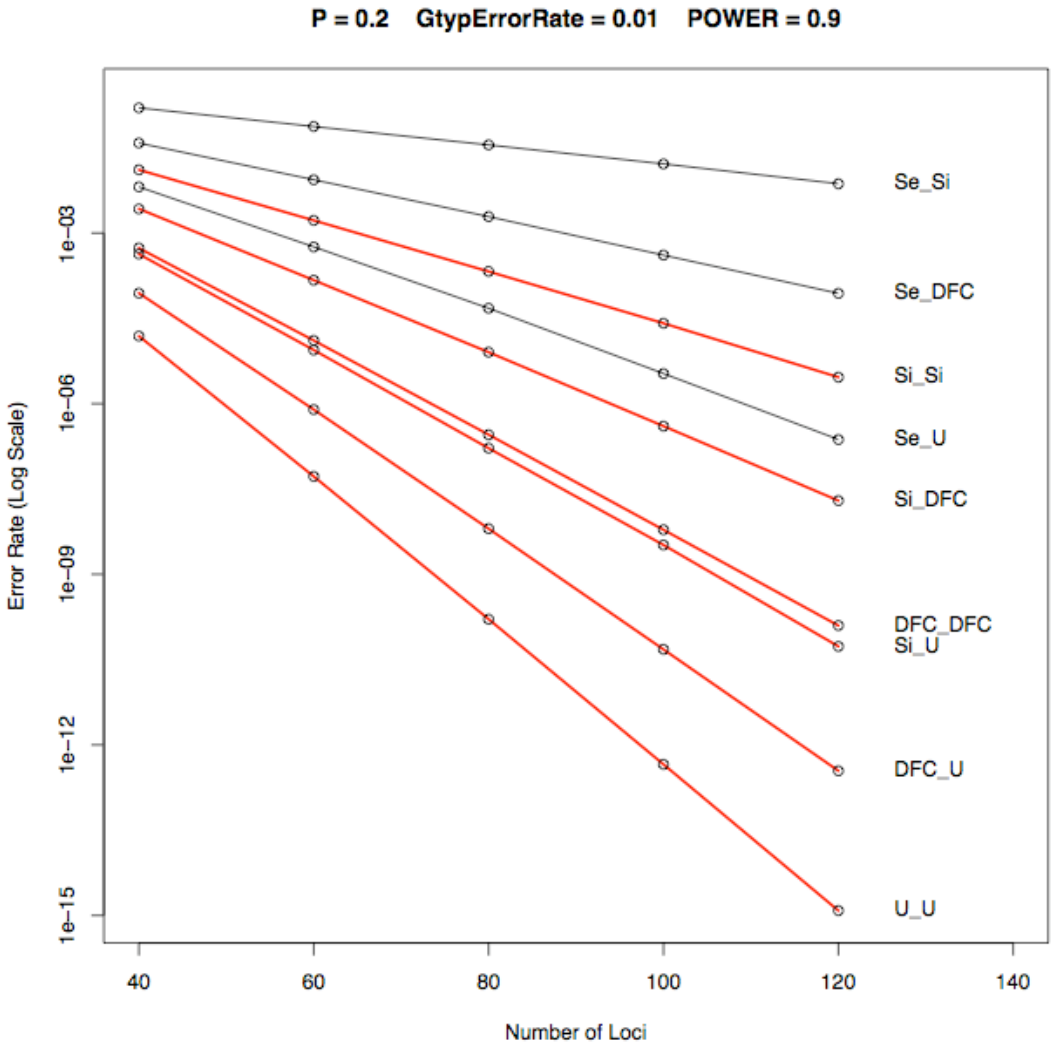


Per-trio false positive rates

80 loci = 4.6×10^{-10}

100 loci = 4.5×10^{-13}

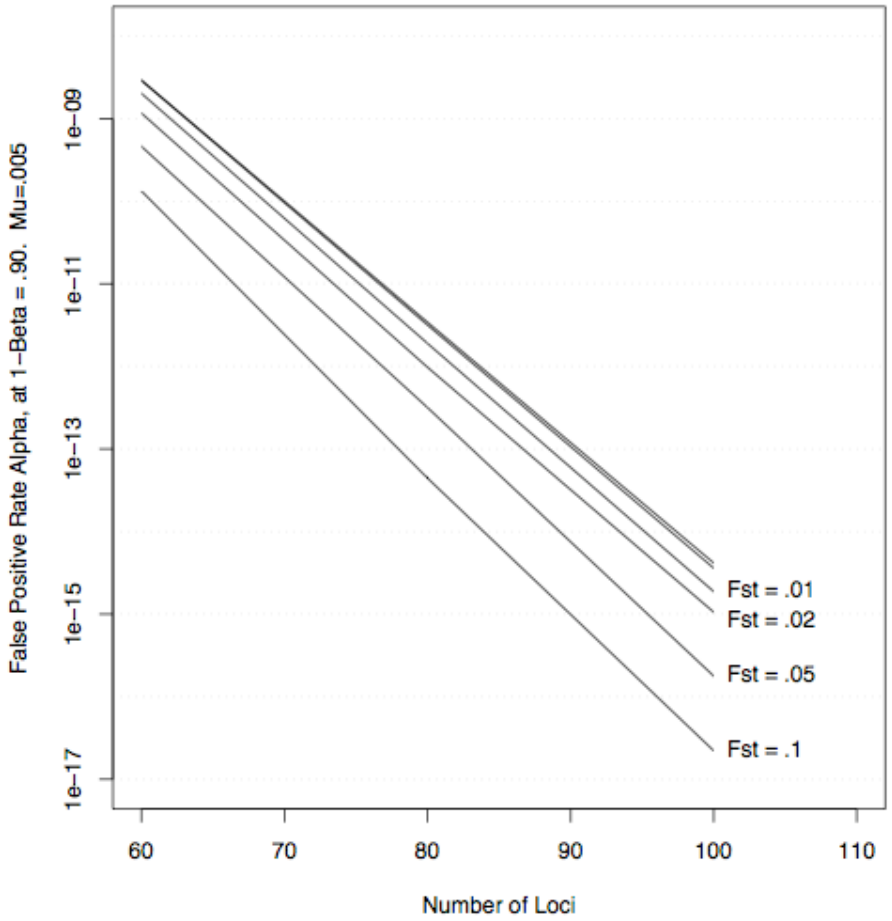
Effect of closely related individuals on error rates



Red: errors which COULD be to wrong hatchery or age

-Close kin can lead to errors, but most common ones don't misidentify either hatchery or cohort.
-Most such combinations are uncommon and/or don't lead to serious errors.
-Recording matings nearly eliminates this problem.

Differentiation makes analyses conservative



-If unrelated individuals in parent database are from differentiated population, then probability of a false positive result with those fish decreases.

-Fst of 0.05 decreases chance of falsely concluding parentage by an order of magnitude.

Advantages of SNPs for large scale parentage-inference

- Low genotyping error rate
- Allele calls (nomenclature) are easily standardized between labs
- Amenable to high-throughput / low-cost genotyping
- Minimal human interaction with the raw genotyping data

Tricks of the trade…or how to reduce the genotyping burden and cost of PBT

- Record matings or sort fish by date of spawning (i.e. day buckets)
- Use only SNP markers with high intermediate allele frequency, p = 0.5 optimal; 10x decrease in error rate with increase of 0.1 in p
- Accept higher error rates in parentage assignment
- Decrease tagging rate (or why its not called FPG anymore)
- Use tag recoveries at escapement in sampling scheme

How much does it cost?

- Cost relative to CWT program is most important because of PST
- Hard to estimate costs of CWT program, but increasing with electronic detection necessary due to mass marking
- Hard to estimate costs of PBT program, as SNP genotyping costs are decreasing.
- Tagging costs for PBT clearly lower than for CWTs
- Tag recovery costs are likely higher for PBT than for CWTs, but…

Costs for SNP genotyping are dropping

FOR IMMEDIATE RELEASE
Tuesday, October 5, 2004

NIH Funds First National SNP Genotyping Center at Broad Institute

Because investigators use different technologies based on the scales and configurations needed, a menu of services will be offered using three different technology platforms. When fully operational, the center will be able to process from 200 million to as many as billions of genotypes per year, depending on the technology platform used and the needs of outside users. The cost for genotyping will be on the order of pennies per genotype, varying according to the technology platform used. Two decades ago, the cost was $10 per genotype, and prices are expected to drop further as technology improves. A portion of the center's annual budget will be used to partially support compelling genotyping research projects, to be selected by a steering committee.

Additional Information from PBT programs

You get much more than stock-of-origin and cohort with PBT

- Reconstruction of large pedigrees
- Map genes for phenotypic traits to locations in the genome
- Near parametric estimates of variance in family size
- Conduct large quantitative genetics studies of phenotypes
- Evaluate different hatchery practices
- Study differences in hatchery and naturally spawning fish by sampling at weirs, fish ladders or carcasses (w/ care)

Parentage-based tagging in fishery management
Future prospects

-Implementation of parentage-based tagging in Central Valley and Klamath Chinook hatcheries to provide cohort of origin for GSI projects in California

-Implementation of integrated PBT and GSI tagging/sampling program for determination of cohort & stock of origin for fish from PBT hatcheries and stock of origin for ALL fish

-Greater understanding of sources of error, costs and limitations of large-scale parentage inference.