

**Genetic Technologies for Monitoring of
Central Valley Chinook
Genetic Stock Identification (GSI)
and
Full Parental Genotyping (FPG)**

Eric C. Anderson* and J. Carlos Garza

*Fisheries Ecology Division
Southwest Fisheries Science Center*

Santa Cruz, CA

August 24, 2005



*A full-blown Linux-Unix freak who supports Open-Source software *and* freely available genetic data

Some Monitoring Objectives for CV Chinook

1. Accurately estimate hatchery contributions to fisheries and spawning populations
2. Accurately estimate straying rates of hatchery fish
3. Estimate ocean survival rates
4. Evaluate the effects of salmonid hatcheries on naturally spawning salmonids
5. Evaluate the effects of hatchery production on genetic composition of populations
6. Estimate rates of hybridization between races of Chinook salmon in hatcheries
7. Evaluate the efficiency of alternative hatchery/release practices
8. Assess homogenization of CV chinook genotypes



Genetic Methods are Uniquely Suited to Identifying the Origin of Individuals, or of Parts of their Genomes

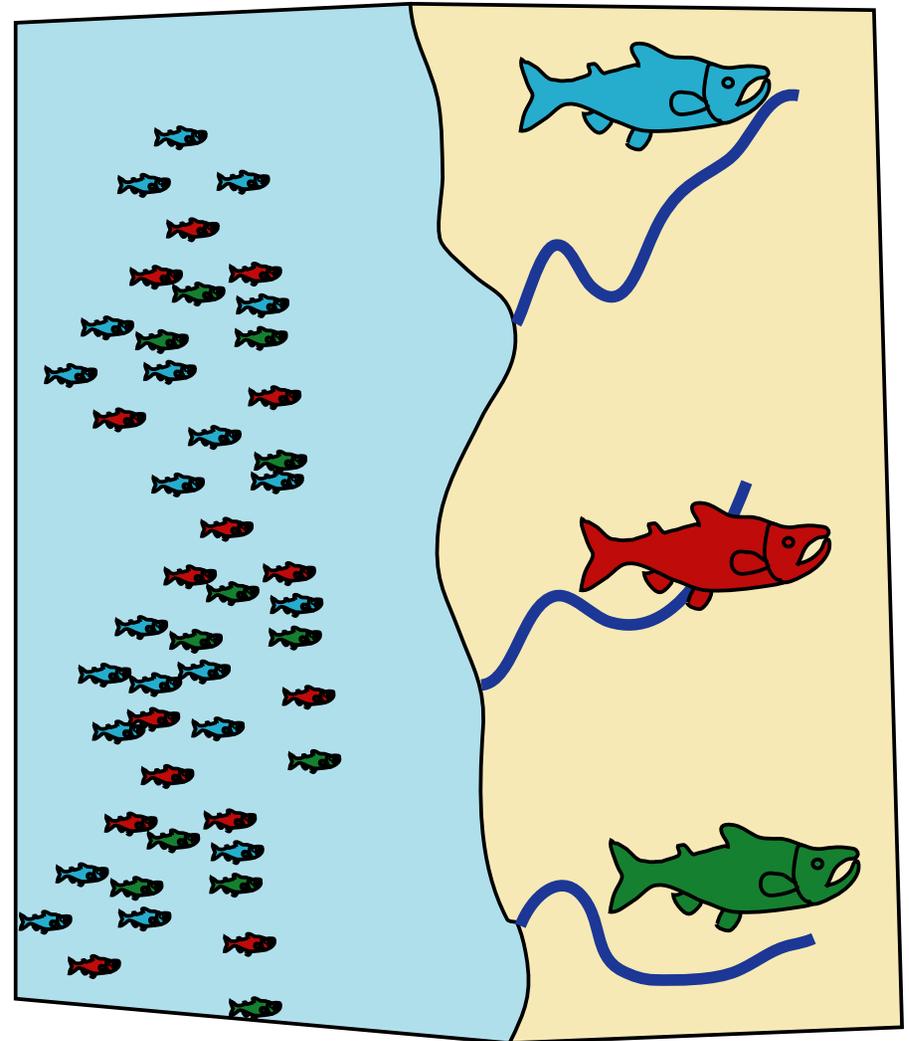
OUTLINE:

- Genetic Stock Identification (GSI)
 - In use for decades
 - Great for identifying the population of origin of individuals
 - Limited utility for very closely related populations
- Full Parental Genotyping (FPG)
 - Emerging technology (on such a large scale)
 - Allows identification of the exact parents of an individual
 - Does not rely on genetic differentiation between populations



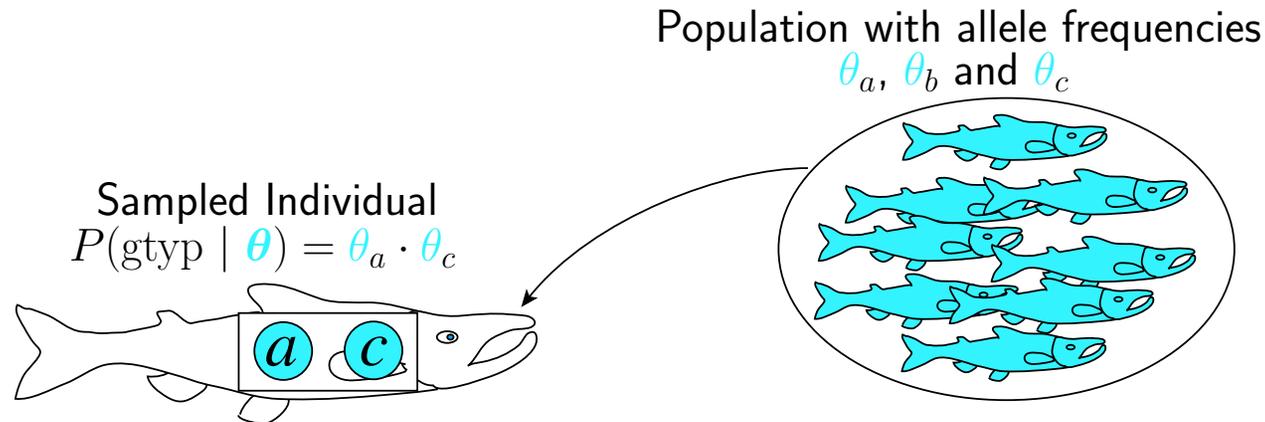
GSI in Cartoon Format

- Our mission is to determine what proportion of fish (being caught) in the ocean are from the red, blue, and green populations.
- Obviously it would be easy if all the fish were color-coded. . . but they are not!
- GSI is a way to use genetic data to determine which population a fish has come from.

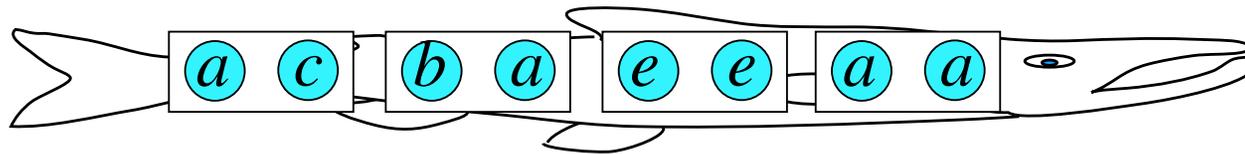


The Model Underlying GSI

- The alleles carried at a locus are independent:
 - Single locus genotype probability



-
- The alleles carried are independent between loci:
 - Multilocus genotype probability



- Today, the loci typically used are *microsatellites*.

An Example—GSI for Klamath *versus* CV Chinook

Background:

- This spring, weak runs of Klamath River Chinook led to dramatically decreased harvest off the California Coast
- This reduction coincided with record returns for CV Chinook.
- Many “surplus” CV chinook went unharvested.
- Losses to the Northern California economy were estimated at \$100 Million

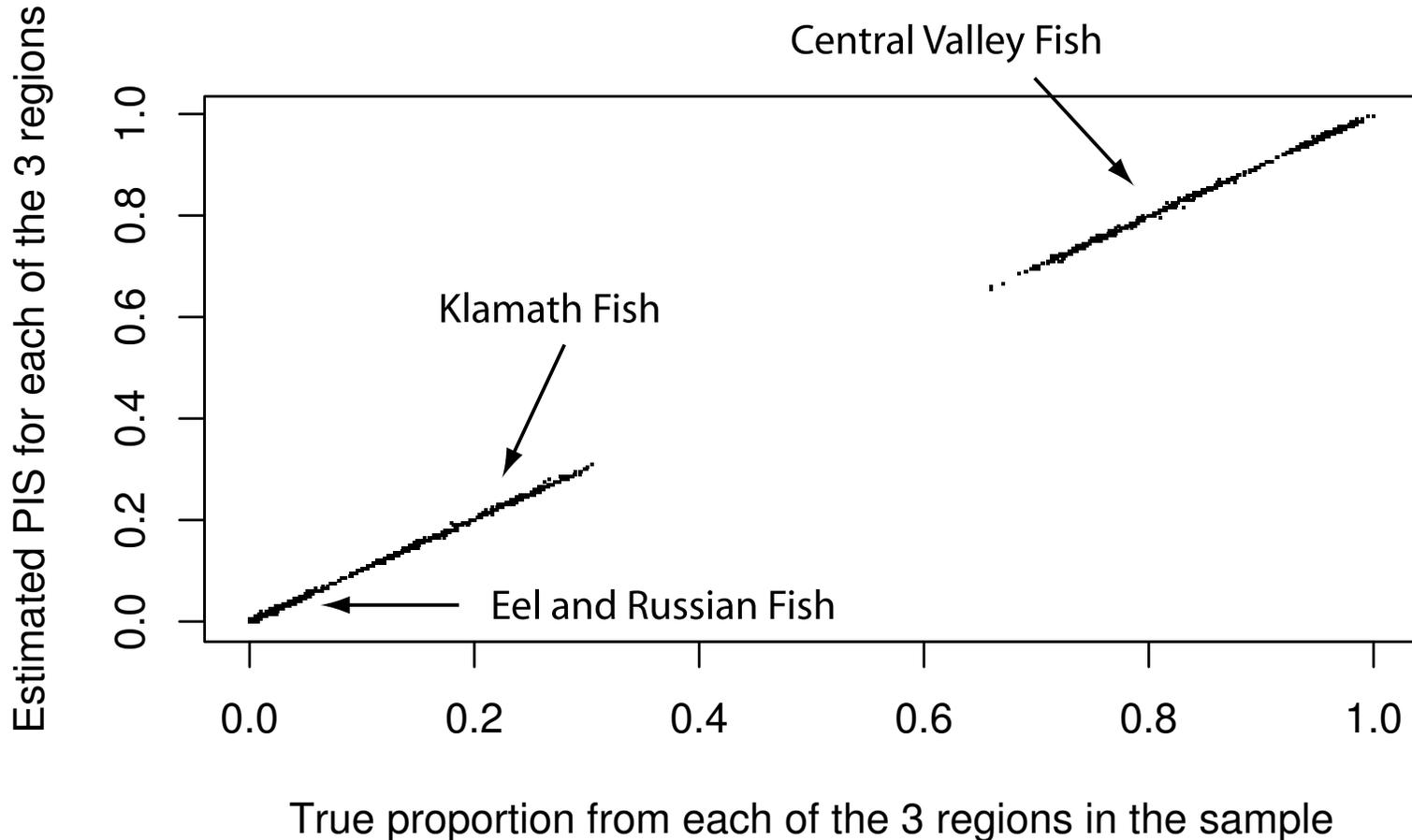
Possible Solution Next Time:

- “Real-Time” GSI
- Better understanding of the ocean distribution of the two stocks



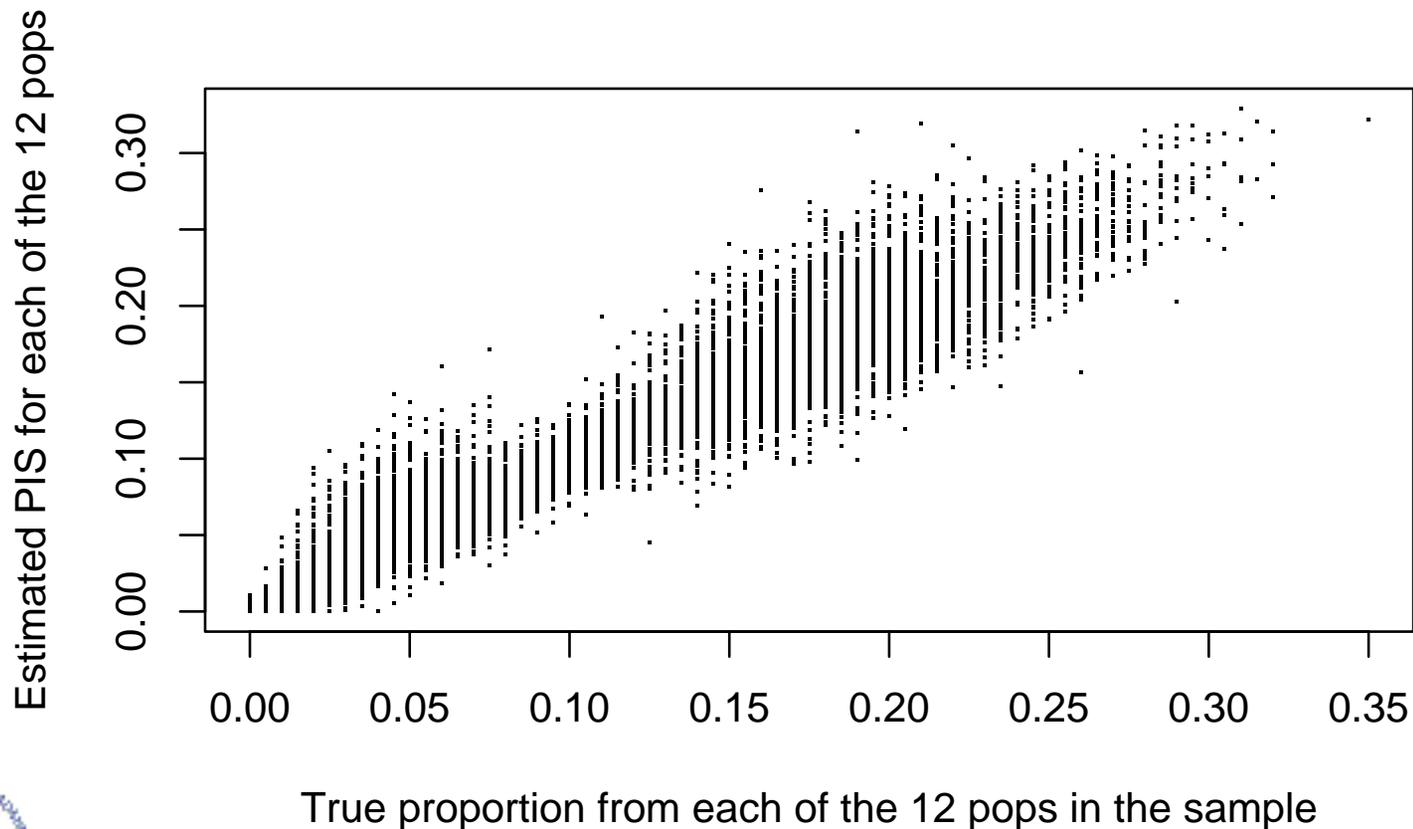
Accuracy of GSI for Estimating Proportions of CV, Coastal, and Klamath Chinook Salmon

— Simulated Samples of Size 200 Fish —



GSI Limitation Number One—Closely Related Populations

Observe What Happens When We Try to Estimate the Proportions of *Specific Populations* (i.e. Battle Creek versus Feather River) in Each Simulated Mixture Sample:



GSI Limitation Number Two—All You Can Get is Population of Origin

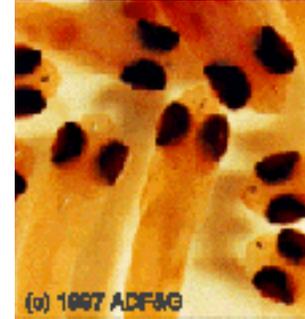
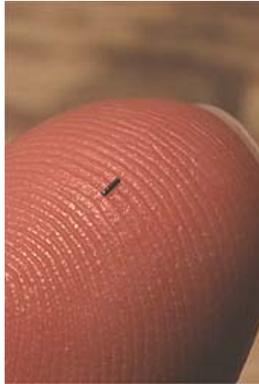
- For many applications in stock assessment and *cohort analysis* it is necessary to know not only the population a fish came from, but also *the year it was born* (its cohort) or the *release group* it was released with.
- Such information is not available with GSI.
- It is typically gleaned through Coded Wire Tagging (CWT)

SEGUE:

- Full Parental Genotyping (FPG) is a genetic tagging method that does yield cohort information (and much, much more!)



Coded Wire Tag Program



Since 1968, 71 agencies in 5 states and B.C. have:

- used 34,000 codes and 573 miles (!!) of wire
- tagged 959,000,000 salmon and steelhead

From 1977 to 1996, only fish with a CWT received adipose fin clips.

Heads of Ad-clipped fish are sent to “head labs” for decoding.

In Alaska, alone, since 1976

- 677 tons of heads (903,000) have been sent to the Juneau head lab.



CWT are recovered from adipose fin-clipped salmon in:

- Terminal, coastal, and high-seas fisheries to estimate fishing rates on particular stocks and cohorts
- Hatchery escapement to estimate survival rates of particular cohorts
- From carcasses recovered from wild spawning grounds
- Hatchery research projects to monitor effects of different hatchery practices

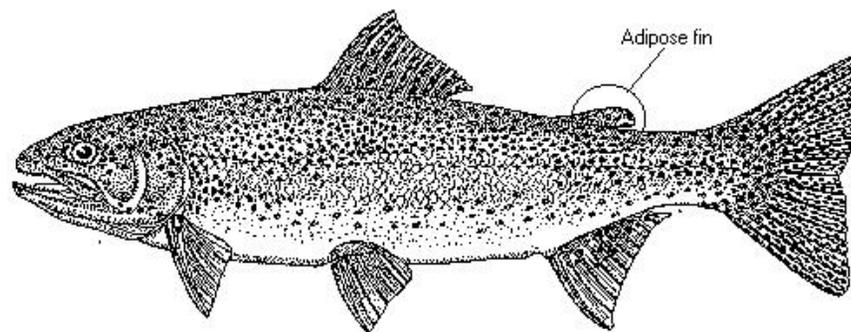
The CWT program has been critical to the Pacific Salmon Commission's goal of estimating (and constraining) fishing pressure on particular stocks and age groups.

The crucial attribute of CWT's is that they allow the identification of the stock of origin, and of the *cohort* of decoded fish.

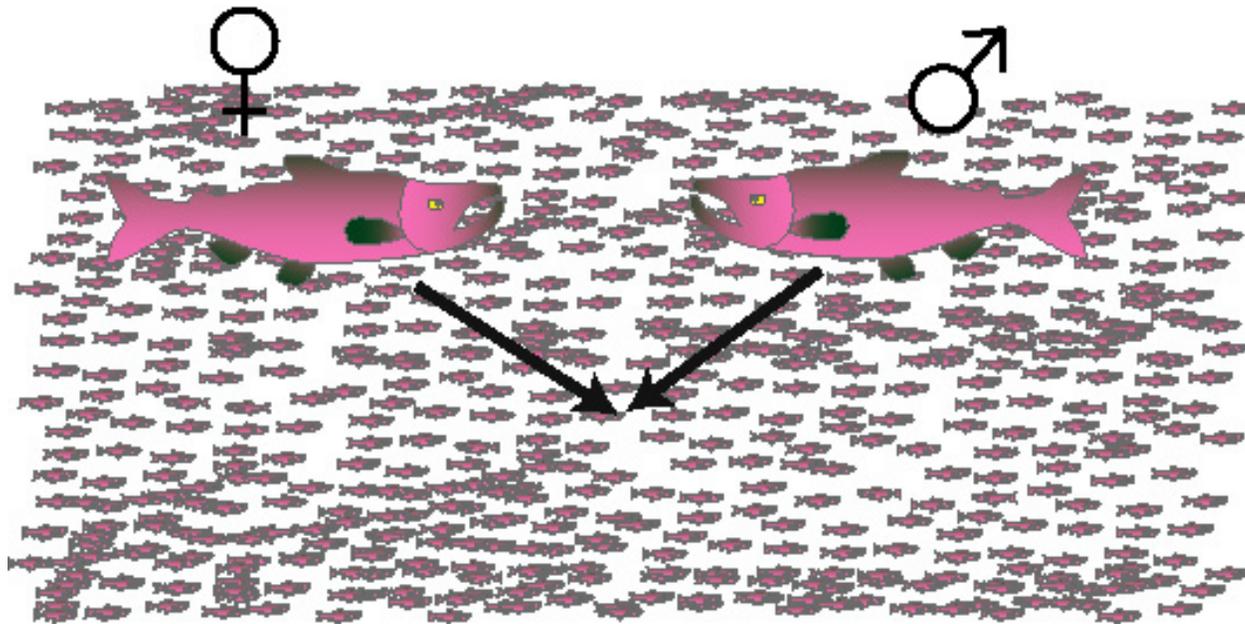


Challenges Facing the CWT Program Today

- 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 for mark-selective fisheries
 - Not all Ad-clipped fish have CWT's



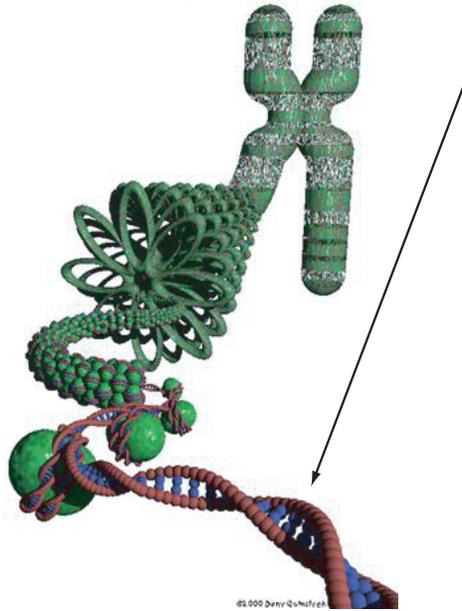
The Logic Behind Full Parental Genotyping



- By genotyping all hatchery parents you can create a database of possible parent pairs and match recovered offspring genotypes against that, assigning offspring to parent pairs (and hence hatchery and cohort).
 - By genotyping two parents, you can effectively tag all their 1,000's of offspring.

SNPs — Single Nucleotide Polymorphisms

...AGGATCGTGCATGGCTAGCTGATTCCGTATCCGTTTC...



Most DNA bases are invariant in a population or species...
But some are variable:

Chromosome 1: ...AGGATCGTGCATGGCT**A**GCTGATTCCGTATCCGTTTC...
Chromosome 2: ...AGGATCGTGCATGGCT**G**GCTGATTCCGTATCCGTTTC...
Chromosome 3: ...AGGATCGTGCATGGCT**A**GCTGATTCCGTATCCGTTTC...
Chromosome 4: ...AGGATCGTGCATGGCT**A**GCTGATTCCGTATCCGTTTC...
Chromosome 5: ...AGGATCGTGCATGGCT**G**GCTGATTCCGTATCCGTTTC...
Chromosome 6: ...AGGATCGTGCATGGCT**G**GCTGATTCCGTATCCGTTTC...
Chromosome 7: ...AGGATCGTGCATGGCT**A**GCTGATTCCGTATCCGTTTC...
Chromosome 8: ...AGGATCGTGCATGGCT**A**GCTGATTCCGTATCCGTTTC...

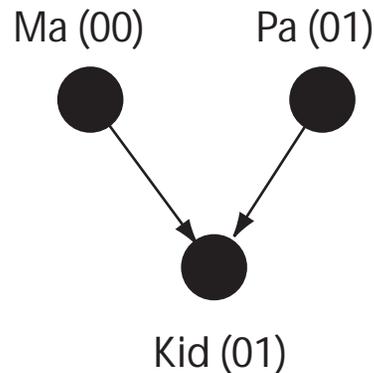
The variant sites are called Single Nucleotide Polymorphisms (SNPS)

- There are usually only two types of bases (letters) found at any SNP.
 - So, call the less frequent base “0” ($\text{freq}=p$) and the more frequent base “1” ($\text{freq}=q = 1 - p$).
 - Each salmon has two copies of the SNP so its *genotype* will be 00 or 01 (or 10) or 11

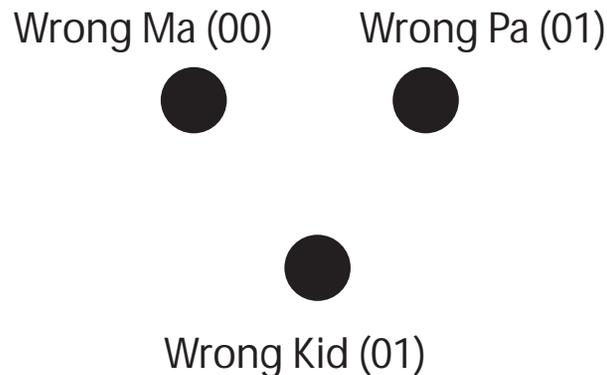


Mendelian Inheritance—The Cornerstone of Parentage Analysis

- Every individual inherits one chromosome from its mother and one from its father.
- This law establishes rules for computing the probability of observing the genotypes of mothers, fathers, and offspring:



$$\Pr(\text{Trio}|\text{Parental}) = p^2 \times 2pq \times \frac{1}{2}$$

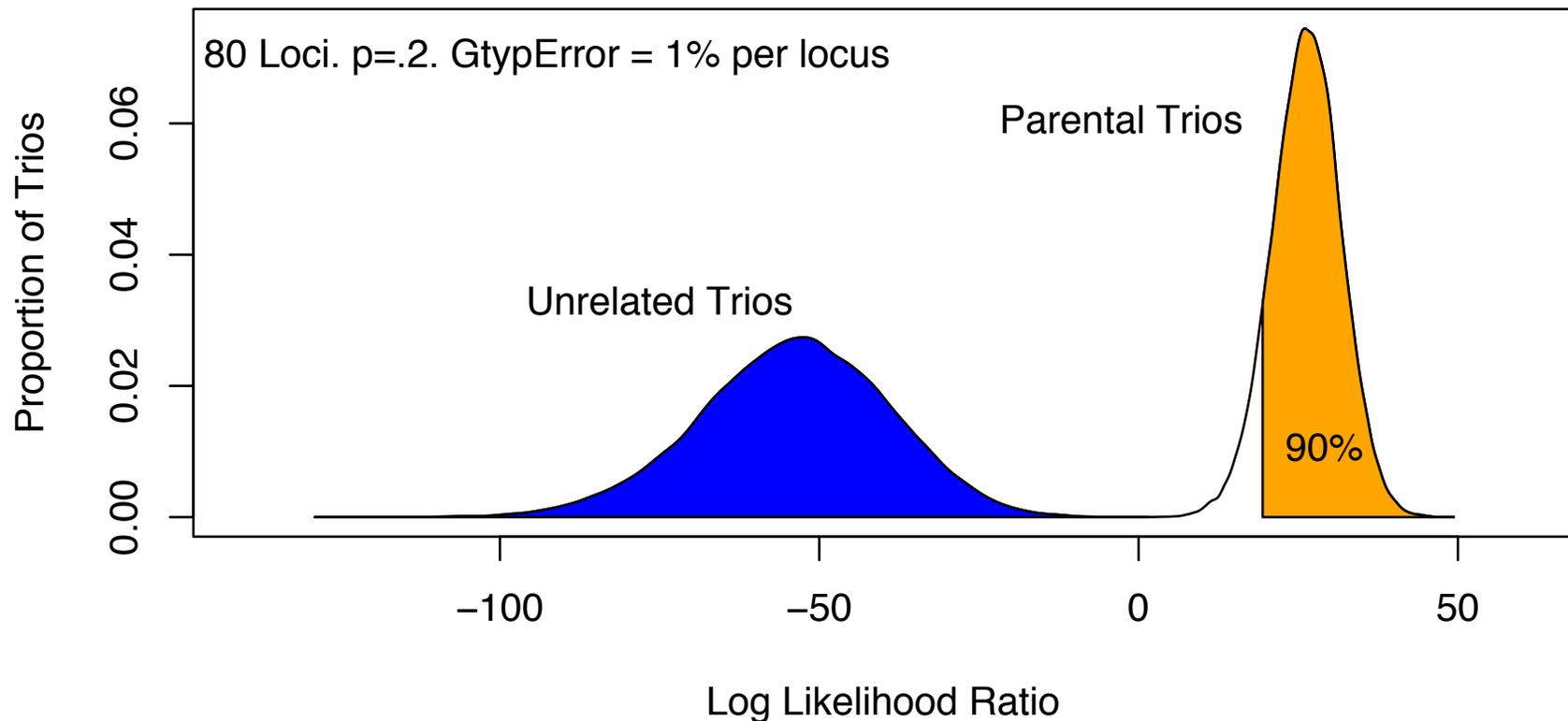


$$\Pr(\text{Trio}|\text{Unrelated}) = p^2 \times 2pq \times 2pq$$



This suggests a natural test statistic—the log likelihood ratio.

90% Power to Correctly Classify True Parental Trios. LogL=19.3



- **Power:** Probability of declaring a true parental trio as parental
 - **False Positive Rate** = Probability of wrongly declaring an unrelated trio as parental

(Gory details in a manuscript submitted to Genetics).



False Positive Rates—Plenty of Chances To Get it Wrong: The number of Parent Pairs in the Database

- CV Fall Chinook Hatchery Example: Assume 30 Million Smolts, 3,000 smolts per female. Same number of males spawned as females, 100 Females and 100 Males Spawned per Day at each hatchery, with spawner carcasses collected in “day-buckets.”
- Number of Females:

$$30 \text{ Million smolts} \times \frac{1 \text{ female}}{3,000 \text{ smolts}} = 10,000 \text{ females}$$

- Number of Parent Pairs:

$$10,000 \text{ females} \times \frac{1 \text{ hatchery} - \text{day}}{100 \text{ females}} \times \frac{10,000 \text{ pairs}}{\text{hatchery} - \text{day}} = 1,000,000 \text{ parent pairs}$$



The Total Number of Trio Comparisons

Let's say the parentage of 50,000 possible offspring will be determined:

$$10^6 \text{ Parent Pairs} \times 50,000 \text{ fish} = 5 \times 10^{10} \text{ trios}$$

Per-Trio False Positive Rates

80 Loci False positive rate of 4.6×10^{-10} per trio

100 Loci False positive rate of 4.5×10^{-13} per trio

So, with as few as 80 loci you could expect to make about 1 mistake in assigning 50,000 fish to their parents.

Total False Positive Rates

80 Loci

$$4.6 \times 10^{-10} \times 5 \times 10^{10} = 23 \text{ out of } 50,000$$



Coded Wire Tag Cost Comparison

Tagging Costs:

- If you can genotype each fish for \$20, you can tag ALL the adipose-fin clipped hatchery fish for the cost of tagging only 10% of them with CWT.
- \$20/fish is a reasonable cost with today's technologies.
- You could get 100% Constant Fractional Marking for one-third of the cost of using CWT's for a 33% CMF scheme.

Recovery Costs:

- The average cost for decoding a recovered CWT in a fish head is about \$5 to \$10 per head.
- So, at today's rates, genotyping each recovered fish would be 2 to 4 times as costly as decoding a CWT in each recovered fish.
- But, some terminal escapement sampling is done as part of tagging . . .
 . . . and genotyping costs are dropping. . .



“Standard” Uses of FPG—Anything You Can Use CWT’s For

- Estimate harvest rates of particular stocks
- Estimate stray rates of different hatchery stocks
- Estimate ocean survival rates
- Estimate incident mortality rates
- *etc.*
- Advantages:
 - Requires only a fin-clip—not a stinky, heavy head.
 - Non-lethal sampling



Additional Uses of FPG

- You get much more than stock-of-origin and cohort
 - Recovering entire pedigrees you can:
1. Estimate variance in family sizes—achieve estimates of effective size of a hatchery population like never before.
 2. Learn about the genetic inheritance of traits like body size and age at maturity
 - Traditional segregation analysis
 - Traditional linkage mapping (for locating genes responsible for traits)
 - Other gene mapping methods like “discordant sib-pair” methods
 - Non-lethal tag recovery for assessing heritage of spring returns
 3. Have huge sample sizes for evaluating different hatchery practices



Concluding Remarks

- GSI and FPG are both genetic methods that offer important solutions for monitoring of CV chinook.
- GSI relies on genetic differentiation between populations, but FPG does not require differentiation.
- FPG offers a potentially cost-effective strategy for tagging all hatchery production
- FPG would offer an enormous amount of data beneficial not only to fisheries managers, but also to hatchery managers and scientists.
- Fish without parents in the database can still be analyzed via GSI.
- We are currently developing the technologies to make FPG possible.

