# Genetic profile for Coleman NFH fall and late fall Chinook salmon

10-29-2010

**Prepared by**

Christian T. Smith

U.S. Fish and Wildlife Service

Abernathy Fish Technology Center

Applied Program in Conservation Genetics

Longview, WA 98632

Christian_Smith@fws.gov, (360) 425-6072 x 337

# National Fish Hatchery Broodstock Genetic Profile

**Facility:** Coleman National Fish Hatchery

**Stock:** Fall Chinook salmon (*Oncorhynchus tshawytscha*)

**Parental stock:** Founding broodstock were collected from Battle Creek and the upper Sacramento River. Since 1987 broodstock have been collected only from Battle Creek.

**Year founded:** Propagated at the Coleman National Fish Hatchery since 1943. Between 1895 and 1942 artificial propagation of fall Chinook had been conducted at the federal Battle Creek Egg Taking Station.

**Generation time:** 2-5 years (y) (10% 2y, 56% 3y, 33% 4y, 1% 5y) [ref 1]

**Segregation / Integration history:** The fall Chinook program at the Coleman NFH is intended to be an integrated program; however, because all hatchery production are not marked, it is not possible to differentiate hatchery and natural fish during spawning and there is no specified target for incorporating natural origin fish. Naturally produced fish are used as broodstock at their rate of collection at the hatchery facility. Fall Chinook in excess of hatchery broodstock needs are managed to fully seed spawning habitats downstream of the hatchery in Battle Creek and spawn naturally.

**Table 1 - Broodstock samples analyzed:**

| Description | Year | n | Life stage | Data source |
|---|---|---|---|---|
| Coleman NFH fall | 2002 | 53 | returning adult | GAPS |
| Coleman NFH fall | 2003 | 60 | returning adult | GAPS |
| Coleman NFH fall | 2005 | 98 | returning adult | AFTC |
| Coleman NFH fall | 2006 | 45 | returning adult | AFTC |

**Facility:** Coleman National Fish Hatchery

**Stock:** Late-fall Chinook salmon (*Oncorhynchus tshawytscha*)

**Parental stock:** Founding broodstock were collected from the upper Sacramento River at the Keswick Dam fish trap. From 1982 to 1996 broodstock were obtained from both Battle Creek (primarily hatchery origin) and the upper Sacramento River (natural origin). From 1997 to 2001 broodstock were collected from only Battle Creek. Since 2002, hatchery broodstock have been collected from Battle Creek and up to 15% natural origin broodstock have been collected from the upper Sacramento River.

**Year founded:** The Coleman National Fish Hatchery late-fall Chinook propagation program was formally separated from the fall Chinook program in 1973, however, since about 1954 fish culturists had recognized differences between these stocks and separated them during spawning.

**Generation time:** 2-5 years (y) (9% 2y, 66% 3y, 24% 4y, 1% 5y) [ref 1]

**Segregation / Integration history:** The late-fall Chinook program at the Coleman National Fish Hatchery is an integrated program, in that attempts are made to incorporate up to 15% naturally produced fish as broodstock on an annual basis. Hatchery origin broodstock are collected from Battle Creek and natural origin broodstock are collected from the upper Sacramento River. Late-fall Chinook originating at the Coleman National Fish Hatchery are not intended to spawn naturally.

**Table 2 - Broodstock samples analyzed:**

| Description | Year | n | Life stage | Data source |
|---|---|---|---|---|
| Coleman NFH late fall | 2005 | 149 | returning adult | AFTC |
| Coleman NFH late fall | 2006 | 34 | returning adult | AFTC |

**Table 3 – Central Valley Chinook salmon samples analyzed for comparison:**

| Description | H/W | Year | n | Life stage | Data source |
|---|---|---|---|---|---|
| Stanislaus River fall | W | 2002 | 65 | returning adult | GAPS |
| Butte Creek fall | W | 2002 | 66 | returning adult | GAPS |
| Butte Creek spring | W | 2002 | 47 | returning adult | GAPS |
| Butte Creek spring | W | 2003 | 68 | returning adult | GAPS |
| Deer Creek spring | W | 2002 | 45 | returning adult | GAPS |
| Feather River H fall | H | 2003 | 136 | returning adult | GAPS |
| Feather River H spring | H | 2003 | 144 | returning adult | GAPS |
| Mill Creek spring | W | 2002 | 62 | returning adult | GAPS |
| Mill Creek spring | W | 2003 | 17 | returning adult | GAPS |
| Sacramento River winter | mix | 1992 | 19 | returning adult | GAPS |
| Sacramento River winter | mix | 1994 | 14 | returning adult | GAPS |
| Sacramento River winter | mix | 1995 | 16 | returning adult | GAPS |
| Sacramento River winter | mix | 1998 | 15 | returning adult | GAPS |
| Sacramento River winter | mix | 2001 | 24 | returning adult | GAPS |
| Sacramento River winter | mix | 2002 | 101 | returning adult | AFTC |
| Sacramento River winter | mix | 2004 | 11 | returning adult | GAPS |
| Sacramento River winter | mix | 2007 | 156 | returning adult | AFTC |
| Sacramento River winter | mix | 2008 | 198 | returning adult | AFTC |

**Genetic markers analyzed:**

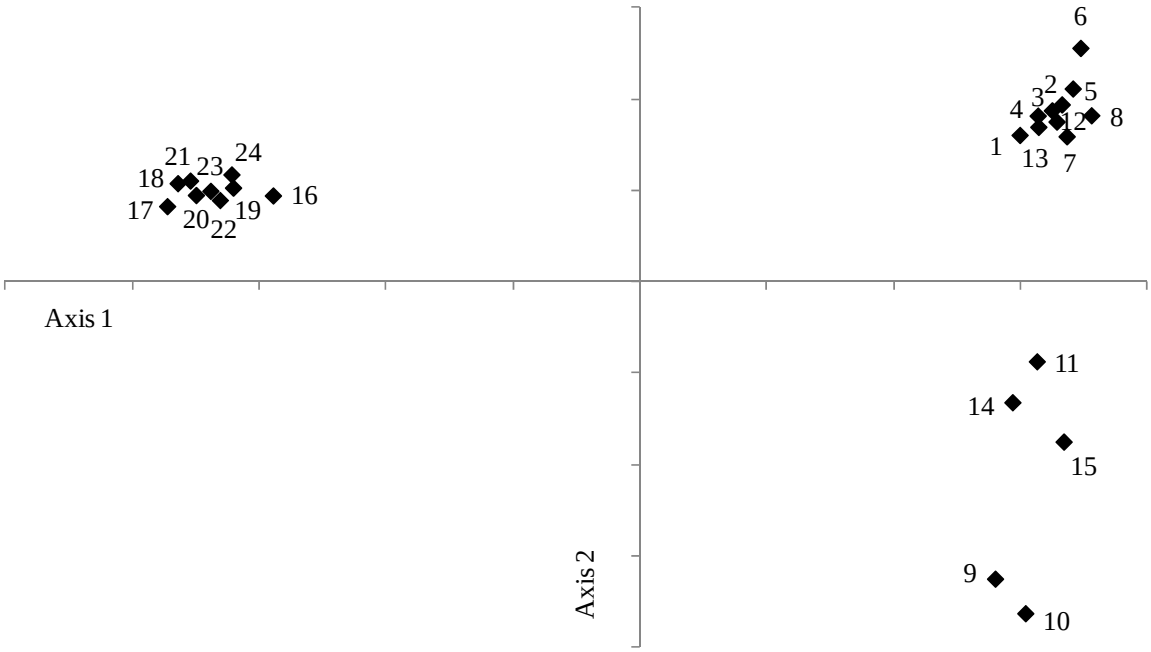
GAPS microsatellites [ref 2] (13 / 13 loci). A complete list of the markers analyzed is provided in Appendix 1. Genotyping success rates are presented in Appendix 2.

**Table 4 - Diversity within samples.** Expected ($H_e$) and observed ($H_o$) heterozygosity, allelic richness rank (AR rank, with 1 representing the most diversity and 21 representing the least diversity), number of loci exhibiting departures from Hardy-Weinberg Equilibrium (HWE), $F_{IS}$ (a measure of departure from random mating), number of pairs of loci exhibiting linkage disequilibrium (LD), and effective population size ($N_e$) based on LD are listed for each sample. Samples in which we observed no evidence for any disequilibrium caused by genetic drift due to a finite number of parents (i.e. those with an estimated $N_e$ of infinity) have $N_e$ marked by "-".

| No | Description | $H_e$ | $H_o$ | AR rank | HWE | $F_{IS}$ | LD | $N_e$ |
|---|---|---|---|---|---|---|---|---|
| 1 | Coleman fall 2002 | 0.85 | 0.85 | 1 | 2 | 0.04 | 0 | - |
| 2 | Coleman fall 2003 | 0.84 | 0.84 | 6 | 3 | 0.05 | 2 | - |
| 3 | Coleman fall 2005 | 0.84 | 0.84 | 5 | 1 | 0.01 | 5 | 2695 (416 - Infinite) |
| 4 | Coleman fall 2006 | 0.84 | 0.84 | 3 | 1 | -0.01 | 1 | 927 (171 - Infinite) |
| 5 | Coleman late fall 2005 | 0.83 | 0.83 | 7 | 5 | 0.01 | 9 | 297 (213 - 466.2) |
| 6 | Coleman late fall 2006 | 0.83 | 0.83 | 11 | 2 | 0.03 | 21 | 106 (59 - 369.4) |
| 7 | Stanislaus_R_02 2002 | 0.85 | 0.85 | 2 | 0 | -0.02 | 2 | - |
| 8 | Butte Creek fall 2002 | 0.83 | 0.83 | 9 | 1 | 0.01 | 5 | 13669 (245 - Infinite) |
| 9 | Butte Creek spring 2002 | 0.79 | 0.79 | 14 | 1 | 0.01 | 2 | - |
| 10 | Butte Creek spring 2003 | 0.79 | 0.79 | 15 | 1 | 0.02 | 1 | - |
| 11 | Deer Creek spring 2002 | 0.83 | 0.83 | 12 | 1 | 0.03 | 4 | 147 (81 - 556.7) |
| 12 | Feather R. H fall 2003 | 0.85 | 0.85 | 4 | 4 | 0.02 | 5 | 1275 (341 - Infinite) |
| 13 | Feather R. H spring 2003 | 0.84 | 0.84 | 8 | 2 | 0.00 | 7 | 898 (362 - Infinite) |
| 14 | Mill Creek spring 2002 | 0.83 | 0.83 | 10 | 4 | 0.01 | 27 | 56 (43 - 75.3) |
| 15 | Mill Creek spring 2003 | 0.82 | 0.82 | 12 | 0 | 0.02 | 1 | - |
| 16 | Sacramento R. winter 1992 | 0.69 | 0.70 | 21 | 1 | -0.06 | 8 | 7 (4 - 10.7) |
| 17 | Sacramento R. winter 1994 | 0.64 | 0.64 | 24 | 1 | -0.08 | 3 | 69 (15 - Infinite) |
| 18 | Sacramento R. winter 1995 | 0.69 | 0.69 | 17 | 0 | -0.04 | 3 | 23 (12 - 81) |
| 19 | Sacramento R. winter 1998 | 0.71 | 0.71 | 20 | 0 | -0.04 | 1 | - |
| 20 | Sacramento R. winter 2001 | 0.68 | 0.68 | 21 | 3 | 0.08 | 3 | 26 (16 - 53.9) |

| No | Description | $H_e$ | $H_o$ | AR rank | HWE | $F_{IS}$ | LD | $N_e$ |
|---|---|---|---|---|---|---|---|---|
| 21 | Sacramento R. winter 2002 | 0.68 | 0.68 | 19 | 1 | 0.01 | 9 | 971 (198 - Infinite) |
| 22 | Sacramento R. winter 2004 | 0.68 | 0.67 | 23 | 0 | 0.01 | 1 | - |
| 23 | Sacramento R. winter 2007 | 0.69 | 0.69 | 18 | 3 | 0.02 | 19 | 279 (177 - 561.6) |
| 24 | Sacramento R. winter 2008 | 0.71 | 0.71 | 16 | 0 | 0.02 | 10 | 1546 (416 - Infinite) |

**Figure 1 - Correspondence Analysis (CA) of allele frequencies observed in Central Valley Chinook salmon populations.** Sample numbers are those listed in Table 3. Axis 1, 2 and 3 accounted for 51.16%, 11.09% and 5.03% of the inertia, respectively.



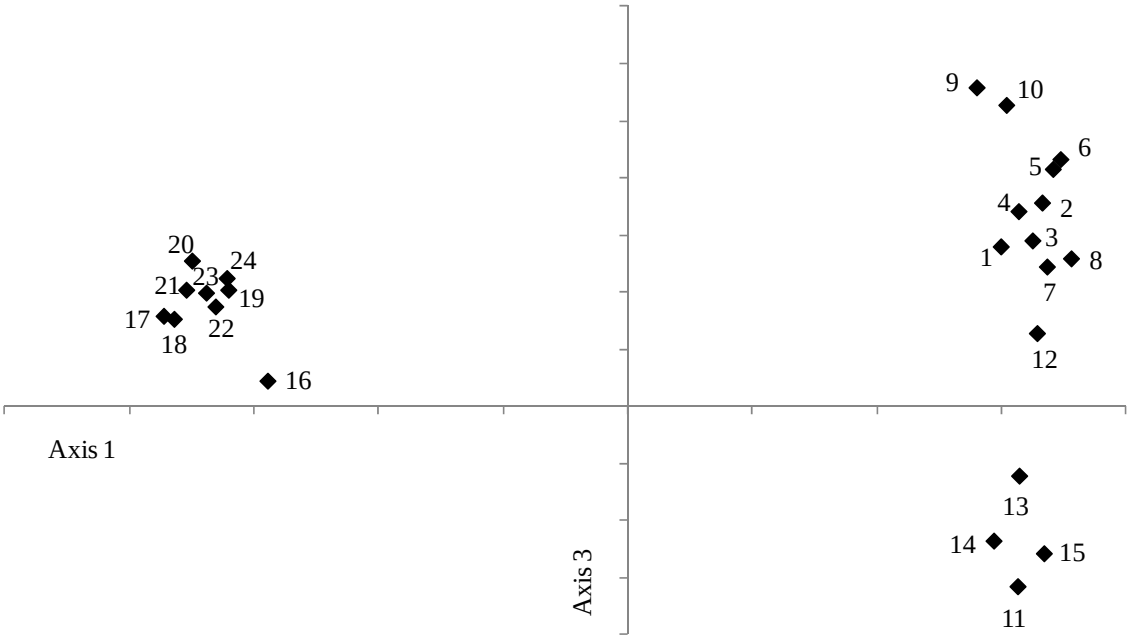
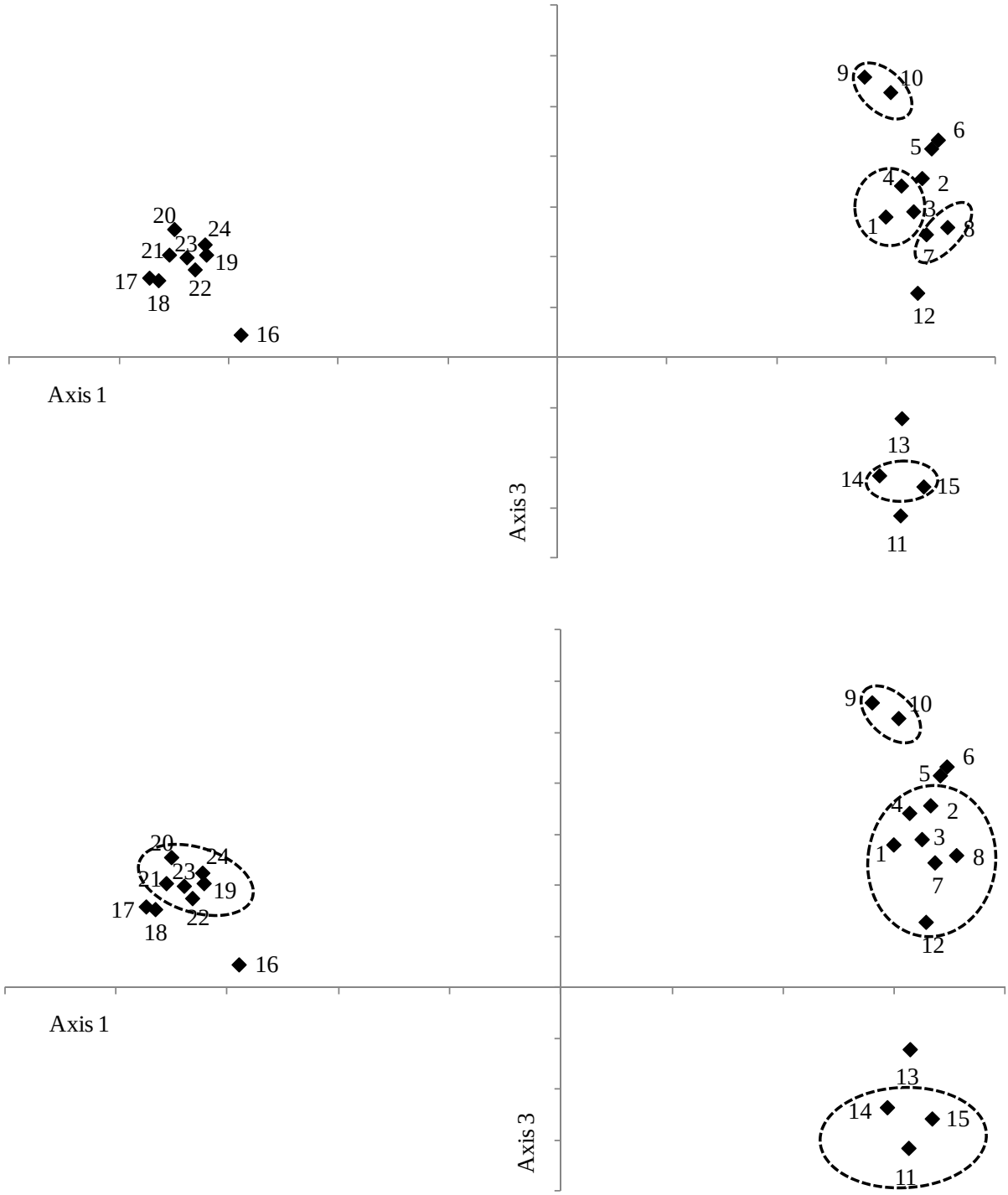



**Figure 2 - Statistical tests of divergence among Central Valley Chinook salmon populations.** Sample numbers are those listed in Table 3. Dashed lines indicate groups of samples lacking statistically different allele frequencies (top) and statistically significant $F_{ST}$ values (bottom).



**Comments**

- The goals of this report series are 1) to summarize available genetic information for NFH broodstocks and make that information available to hatchery managers, and 2) to make sure that data for the NFH broodstocks are available for internal hatchery reviews and HET meetings, as well as to our partners.
- Genetic diversity (measured by heterozygosity and allelic richness) appeared high in Coleman NFH fall Chinook salmon relative to all other Central Valley stocks examined. Genetic diversity in Coleman NFH late fall Chinook salmon was somewhat lower, but was high relative to winter run or the spring run stocks examined.
- Estimates of LD and Ne indicate that the effective size of the Coleman NFH fall Chinook salmon population is very large. We detected little evidence of increased inbreeding (relative to a population of infinite size). Ne for the late fall run was small compared to Coleman NFH fall run or to recent years of winter run. Substantial LD observed in the late fall samples may indicate family structure (a group of closely related individuals in the sample), or could also reflect a run mixture (i.e.: fall and late fall individuals) in that sample.
- Divergence was low among annual samples of Coleman NFH fall run and other Central Valley fall run samples, as indicated by non-significant $F_{ST}$ values (Figure 2). In contrast, $F_{ST}$ between the two late fall samples was significant. Significant $F_{ST}$ between years could reflect family structure, a run mixture in one or both samples, or genetic drift among years. Samples from additional years would be helpful for interpreting this result.

## Sources cited

1. Red Bluff Fish and Wildlife Office. 2010. Unpublished Data

2. Information regarding the microsatellite markers used and the inter-agency baseline in which they are used may be found in the following article:

Seeb LW, Antonovich A, Banks MA, Beacham TD, Bellinger MR, Campbell M, Garza JC, Guthrie CM III, Moran P, Narum SR, Stephenson JJ, Supernault KJ, Teel DJ, Templin WD, Wenburg JK, Young SF, Smith CT. 2007. Development of a Standardized DNA Database for Chinook salmon. Fisheries 32: 540 - 552.

## Availability

All data generated for this report will be included in the next builds of the shared GAPS consortium baselines (hosted by NOAA Fisheries). Allele frequencies for NFH stocks are also available from Abernathy Fish Technology Center upon request.

## Disclaimer

The findings and conclusions in this report are those of the author and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

**Appendix 1.** Thirteen microsatellite markers used to analyze Coleman NFH fall and late fall Chinook salmon.

| Microsatellite Markers | |
|---|---|
| 1 | Ogo2 |
| 2 | Ogo4 |
| 3 | Oki100 |
| 4 | Omm1080 |
| 5 | Ots201b |
| 6 | Ots208b |
| 7 | Ots211 |
| 8 | Ots212 |
| 9 | Ots213 |
| 10 | Ots3M |
| 11 | Ots9 |
| 12 | OtsG474 |
| 13 | Ssa408 |

**Appendix 2.** Genotype call rates (completeness of data). Shaded cells indicate data that are <90% complete.

| Description | Year | Microsatellite call rate |
|---|---|---|
| Coleman NFH broodstock | 2002 | 0.922 |
| Coleman NFH broodstock | 2003 | 0.926 |
| Coleman NFH broodstock | 2005 | 0.997 |
| Coleman NFH broodstock | 2006 | 0.985 |
| | | |
| Battle Creek late fall | 2005 | 0.989 |
| Battle Creek late fall | 2006 | 0.977 |
| Stanislaus_R_02 | 2002 | 0.928 |
| Butte Creek fall | 2002 | 0.945 |
| Butte Creek spring | 2002 | 0.908 |
| Butte Creek spring | 2003 | 0.916 |
| Deer Creek spring | 2002 | 0.952 |
| Feather River H fall | 2003 | 0.931 |
| Feather River H spring | 2003 | 0.986 |
| Mill Creek spring | 2002 | 0.955 |
| Mill Creek spring | 2003 | 0.964 |
| Sacramento River winter | 1992 | 0.951 |
| Sacramento River winter | 1994 | 0.912 |
| Sacramento River winter | 1995 | 0.923 |
| Sacramento River winter | 1998 | 0.872 |
| Sacramento River winter | 2001 | 0.955 |
| Sacramento River winter | 2002 | 0.997 |
| Sacramento River winter | 2004 | 0.895 |
| Sacramento River winter | 2007 | 0.999 |
| Sacramento River winter | 2008 | 0.999 |

**Appendix 3.** Pairwise $F_{ST}$ values between collections of Central Valley Chinook salmon. Sample numbers are those listed in Table 3. Shaded cells indicate non-significant results (pairwise $F_{ST}$ values not >95% of a null distribution in which individuals were permuted among samples).

| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.000 | -0.001 | -0.002 | 0.003 | 0.009 | 0.002 | 0.003 | 0.026 | 0.024 | 0.015 | 0.001 | 0.006 | 0.012 | 0.013 | 0.109 | 0.145 | 0.126 | 0.111 | 0.129 | 0.132 | 0.113 | 0.126 | 0.119 |
| 2 | - | 0.001 | 0.001 | 0.004 | 0.009 | 0.002 | 0.002 | 0.029 | 0.028 | 0.014 | 0.003 | 0.010 | 0.015 | 0.016 | 0.119 | 0.153 | 0.137 | 0.121 | 0.139 | 0.140 | 0.128 | 0.133 | 0.127 |
| 3 | | - | 0.000 | 0.005 | 0.012 | 0.004 | 0.005 | 0.033 | 0.030 | 0.016 | 0.004 | 0.009 | 0.017 | 0.016 | 0.114 | 0.152 | 0.133 | 0.120 | 0.137 | 0.134 | 0.123 | 0.129 | 0.123 |
| 4 | | | - | 0.004 | 0.012 | 0.003 | 0.003 | 0.028 | 0.026 | 0.017 | 0.004 | 0.011 | 0.017 | 0.020 | 0.118 | 0.159 | 0.134 | 0.118 | 0.138 | 0.141 | 0.126 | 0.134 | 0.127 |
| 5 | | | | - | 0.002 | 0.005 | 0.003 | 0.036 | 0.033 | 0.018 | 0.005 | 0.013 | 0.020 | 0.016 | 0.123 | 0.161 | 0.141 | 0.128 | 0.146 | 0.145 | 0.137 | 0.138 | 0.133 |
| 6 | | | | | - | 0.009 | 0.005 | 0.048 | 0.044 | 0.025 | 0.009 | 0.017 | 0.027 | 0.026 | 0.133 | 0.171 | 0.149 | 0.137 | 0.160 | 0.159 | 0.144 | 0.150 | 0.144 |
| 7 | | | | | | - | 0.002 | 0.026 | 0.024 | 0.014 | 0.002 | 0.006 | 0.012 | 0.013 | 0.112 | 0.149 | 0.132 | 0.117 | 0.135 | 0.138 | 0.126 | 0.132 | 0.126 |
| 8 | | | | | | | - | 0.036 | 0.031 | 0.015 | 0.002 | 0.010 | 0.018 | 0.016 | 0.127 | 0.167 | 0.148 | 0.135 | 0.153 | 0.153 | 0.142 | 0.147 | 0.141 |
| 9 | | | | | | | | - | -0.002 | 0.026 | 0.029 | 0.031 | 0.022 | 0.026 | 0.125 | 0.161 | 0.157 | 0.125 | 0.144 | 0.151 | 0.140 | 0.142 | 0.136 |
| 10 | | | | | | | | | - | 0.023 | 0.027 | 0.029 | 0.017 | 0.021 | 0.127 | 0.165 | 0.158 | 0.132 | 0.149 | 0.154 | 0.141 | 0.146 | 0.140 |
| 11 | | | | | | | | | | - | 0.014 | 0.015 | 0.003 | 0.003 | 0.119 | 0.157 | 0.146 | 0.128 | 0.147 | 0.148 | 0.135 | 0.142 | 0.136 |
| 12 | | | | | | | | | | | - | 0.002 | 0.014 | 0.013 | 0.109 | 0.147 | 0.131 | 0.117 | 0.135 | 0.133 | 0.123 | 0.128 | 0.123 |
| 13 | | | | | | | | | | | | - | 0.014 | 0.015 | 0.105 | 0.144 | 0.133 | 0.117 | 0.134 | 0.130 | 0.120 | 0.127 | 0.121 |
| 14 | | | | | | | | | | | | | - | 0.003 | 0.116 | 0.153 | 0.138 | 0.125 | 0.142 | 0.140 | 0.125 | 0.133 | 0.129 |
| 15 | | | | | | | | | | | | | | - | 0.130 | 0.178 | 0.155 | 0.138 | 0.162 | 0.160 | 0.147 | 0.149 | 0.143 |
| 16 | | | | | | | | | | | | | | | - | 0.054 | 0.046 | 0.033 | 0.041 | 0.039 | 0.050 | 0.036 | 0.036 |
| 17 | | | | | | | | | | | | | | | | - | 0.054 | 0.036 | 0.024 | 0.018 | 0.035 | 0.019 | 0.027 |
| 18 | | | | | | | | | | | | | | | | | - | 0.027 | 0.017 | 0.015 | 0.032 | 0.011 | 0.011 |
| 19 | | | | | | | | | | | | | | | | | | - | 0.003 | 0.009 | 0.016 | 0.009 | 0.008 |
| 20 | | | | | | | | | | | | | | | | | | | - | 0.008 | 0.017 | 0.011 | 0.006 |
| 21 | | | | | | | | | | | | | | | | | | | | - | 0.008 | 0.001 | 0.001 |
| 22 | | | | | | | | | | | | | | | | | | | | | - | 0.009 | 0.012 |

**Appendix 4.** Glossary

**Allele** – A unique genetic character state. Each locus has two alleles.

**Allelic richness** – The number of alleles observed in a sample of individuals, corrected for unequal sample sizes by rarefaction.

**Effective population size ($N_e$)** – The number of individuals in a model population which would lose genetic variation at the same rate as an observed population. Deviations from model behavior in real populations (e.g. unequal sex ratios, some individuals reproducing more than others, etc…) tend to make $N_e$ lower than census size (N).

**$F_{IS}$** – Correlation of alleles in an individual relative to the subpopulation in which it occurs. Commonly used as a measure of departure from random mating within a subpopulation.

**$F_{ST}$** – Correlation of alleles within the same subpopulation relative to the entire population. Commonly used as a measure of divergence between subpopulations.

**Gene flow** – Movement of genetic material from one population to another. Implies both physical movement and successful integration into the recipient population.

**Linkage Disequilibrium (LD)** – A measure of departure from independence of alleles in a pair of loci.

**Genetic Drift** – Process of genetic divergence between populations based on random sampling of alleles each generation.

**Heterozygosity** – Proportion of individuals in a population that are heterozygotes (i.e. do not have two identical alleles at a locus).

**Hardy-Weinberg Equilibrium (HWE)** – Genotype ratios expected under a random mating model.

**Locus** – A physical location on the DNA of an organism. The term "locus" is often used synonymously with "marker" or with any type of marker (e.g., "SNP" or "microsatellite").