Summary of hatchery-wild introgressive hybridization for northern Puget Sound steelhead (*Oncorhynchus mykiss*) populations affected by segregated hatchery programs

Kenneth I Warheit

Washington Department of Fish and Wildlife
Molecular Genetics Laboratory
600 Capitol Way N.
Olympia, WA 90501-1091
kenneth.warheit@dfw.wa.gov

Unpublished report.
Washington Department of Fish and Wildlife
March 2014

Introduction

Fish and wildlife managers have been translocating and artificially propagating animals for decades, generally for the purpose of conserving declining or depressed populations or species, or for supplementing a population that is exploited for commercial or recreational harvest (Frankham et al. 2002, Naish et al. 2008, Laikre et al. 2010). The use of hatcheries to enhance salmon and steelhead populations has been extensive, and our understanding of their effects on wild populations is growing (e.g., Waples 1991, Hilborn 1992, Araki et al. 2007, Laikre et al. 2010, Thériault et al. 2011, Christie et al. 2012, Hess et al. 2012, Seamons et al. 2012, Zhivotovsky et al. 2012). These effects can be categorized as either ecological (e.g., competition), or genetic (e.g., domestication, hybridization) (Naish et al. 2008, Kostow 2009). To moderate or eliminate the negative effects to wild populations the Hatchery Scientific Review Group (HSRG) recommended that every salmonid hatchery develop a genetic management plan, and every hatchery population be managed as either segregated from or integrated with the wild population(s) that spawn naturally within the same basin (HSRG 2004, Mobrand et al. 2005).

The intent of segregated hatchery programs is to keep separate the hatchery and wild populations, and they are managed so that only hatchery-origin individuals are used as broodstock, and hatchery-origin adults are restricted from spawning naturally, with the understanding that natural spawning by hatchery-origin fish from the segregated program will impose potential risks to natural populations. Therefore, by design the hatchery and wild populations in segregated programs are genetically distinct, and the degree of genetic differentiation is a function of the source of the hatchery broodstock, hatchery founder effect, genetic drift, or domestication selection (Mobrand et al. 2005). Primarily, the purpose of segregated hatchery programs is to create harvest opportunities, and secondarily, to direct harvest away from wild populations of conservation concern (e.g., mark-selective fishery). However, if segregated hatchery-origin individuals return as adults and stray away from their hatchery of origin, an unintended consequence of a segregated hatchery program would be hybridization between hatchery-origin and wild fish that spawn naturally. Hybridization may be unavoidable if fishery managers lack the ability to restrict hatchery-origin fish from natural spawning grounds, and if spawning by hatchery-origin and wild fish is not segregated spatially or temporally (Waples 1991, Naish et al. 2008). One of the challenges in managing a segregated hatchery program is to monitor the number of hatchery-origin fish that stray into natural spawning areas, and to document the degree to which these fish interact reproductively with wild fish. This is especially true for species that are difficult to observe and monitor while they spawn, such as steelhead trout (*Oncorhynchus mykiss*).

The main purpose of this document is two-fold: (1) to provide a summary of the genetic composition of steelhead populations within the Green, Snohomish, Stillaguamish, Skagit, Samish, and Nooksack River basins, and (2) to discuss in detail the methods used to obtain these genetic compositions. This document is not intended to be a comprehensive assessment of the genetic diversity and differentiations of Puget Sound steelhead, nor is it a treatment as to the hatchery-based causes of introgressive hybridization. This document is also an addendum to each of the Hatchery Genetic Management Plans (HGMPs; 2014 versions) for Washington Department of Fish and Wildlife steelhead hatchery facilities in the Green, Snohomish, Stillaguamish, Skagit, and Nooksack River basins.

This document is organized mainly as Methods and Results, with minimal discussion within the Results, and with a short Conclusion section. To establish measures of hatchery-wild introgression we used the program *Structure* (Pritchard et al. 2000, Falush et al. 2003) on empirical data to assign individuals to categories or groups of individuals. We also created modeled populations that simulated the empirical data, and then ran *Structure* on the modeled populations to estimate *Structure's* assignment error. We adjusted or corrected *Structure's* assignments of the empirical data to account for these errors. Finally, we characterized the hatcheries' genetic effects on wild populations using two measures: effective pHOS, and introgression. We identified hatchery-wild introgression in all river basins, but the level of introgression varied both within and among basins. We preliminarily explored some of the potential reasons for this variability.

Methods

In this Methods Section we outline a series of steps involving the formation and analysis of modeled populations and analyses of wild and hatchery steelhead samples (Figure 1), to produce measures of hatchery-wild introgressive hybridization. We used two primary analysis tools: (1) the program *Structure* (Pritchard et al. 2000, Falush et al. 2003), and (2) a likelihood-based procedure that adjusts or corrects *Structure* results to account for the close phylogenetic relationships (i.e., recency of common ancestry) between the hatchery and wild populations (Warheit and Knapp, in prep). The modeled populations provide data to objectively establish thresholds to assign individuals to *Structure* groups or categories, and measures of assignment errors. We include the model-based assignment errors into the likelihood procedure to produce final estimates of population composition and hatchery-wild introgression.

Samples

All samples used in this analysis were fin tissue samples archived in the Washington Department of Fish and Wildlife Molecular Genetics Laboratory (WDFW-MGL) tissue collection. Each sample collection is accessioned with a WDFW-MGL code, and most collections are associated with field collection data that included collection year and location, age or life stage of individual samples, collection dates, origin (hatchery [adipose fin absent] versus natural [adipose fin present]), and a presumed run timing (Tables 1a, 2). We refer to natural-origin fish as wild, although these wild fish may have hatchery- and /or natural-origin ancestors. Our focus here was steelhead wild and hatchery collections in north Puget Sound, so we limited collections for genotyping and analysis to those located in the Green, Snohomish, Stillaguamish, Skagit, Samish, and Nooksack River basins. We combined wild samples with similar collection year and dates, life stage, origin, and presumed run timing into collection aggregates, which we called Operational Units (OUs) (Table 1b). Operational Units were the primary unit for analyses. Operational Units were combined into NOAA PSSTRT designated Demographically Independent Populations (DIPs) (PSSTRT 2013), based on their location and presumed run timing (Table 1b), and were also subjected to statistical analyses. All hatchery collections (Table 2) were limited to steelhead segregated programs, which in Puget Sound include early winter hatchery (EWH) programs, which were derived initially from Chambers Creek, Puget Sound, Washington, and early summer hatchery (ESH) programs, which were derived initially from Skamania Hatchery, Washougal River, lower Columbia, Washington.

Genotypes

We used two 96 single nucleotide polymorphism (SNP) panels, for a total of 192 SNPs, to genotype all samples (Table 3). These panels, designated by WDFW-MGL as Panels E and F, were designed for the purpose of genotyping *Oncorhynchus mykiss* (steelhead, and rainbow and redband trout) samples throughout Washington State as baseline samples for genetic stock identification (GSI), population differentiation, and hatchery management, and not specifically for measuring hatchery-wild introgression. We included as part of these panels three SNPs designed to identify pure cutthroat trout (*O. clarki*), or cutthroat – steelhead hybrids, both of which occur in north Puget Sound river systems, and could be confused with pure steelhead samples, especially juvenile samples. We identified every SNP by a laboratory-designated name, but provide the official locus name and source reference in Table 3.

Sample and genotype quality assurance (QA)

For the entire dataset, we removed loci if there was no variation across all individuals, or if fewer than 80% of the individuals were scored. For individual basin analyses we removed a locus if it was not scored for an entire OU. After surveying for cutthroat trout alleles, we removed the three loci designed to identify pure cutthroat trout or cutthroat – steelhead hybrids. Samples were removed for three reasons: (1) if more than one-third ($N \geq 63$) of the loci were missing; (2) if the sample was scored with one of more cutthroat alleles at any one of three designated cutthroat specific loci; or (3) if the sample showed a relatedness (r) of half-sib or closer to at least one other sample in the OU. We used the program ML-Relate (Kalinowski et al. 2006) to measure the pairwise relatedness between individuals within an OU. On a pairwise basis, we removed one individual of the pair with a relatedness coefficient of 0.25 or more, and repeated the process until no pair of individuals within the dataset appeared related at the half-sibling level or closer ($r \geq 0.25$; see Blouin 2003).

Constructing model populations

In order to establish objective assignment thresholds in the empirical data, and to estimate assignment error, we constructed three model populations using the program *MS* (Hudson 2002), which builds coalescent trees that approximate evolution under a Wright-Fisher model (*sensu* Crow and Kimura 1970) based on a series of user-defined parameters. The intent of these modeled populations was to simulate the empirical genetic structure among a wild, an early winter hatchery (EWH), and an early summer hatchery (ESH) populations within a given river basin. Our first requirement of these populations was that they were related hierarchically (Figure 2), with the more distantly related population (Pop 3 in Figure 2) being the ESH populations, and Pop 1 and Pop 2 (in Figure 2) arbitrarily assigned to wild and EWH populations, respectively. Our second requirement of the populations was that their genetic distances should model closely the genetic relationships among the wild, EWH, and ESH OUs within each basin. We used F_{ST} as our measure of genetic distance, and for each basin we calculated three F_{ST} values, averaged across all OUs within the basin: (1) wild (summer or winter) versus within-basin early winter hatchery (or surrogate, if EWH not present in basin), (2) wild (summer or winter) versus within-basin early summer hatchery (or surrogate, if ESH not present in basin), and (3) early winter hatchery versus early summer hatchery (or surrogate, if neither or both EWH or ESH were not present in basin) (Table 4). We used Cockerham's formulation of F_{ST} (Cockerham 1969), calculated using a *Matlab* (MathWorks 2012) custom

script. For our initial estimate of pairwise F_{ST}s for each basin, we used the final N for each OU, after samples were removed (Tables 1b, 2), as explained above, and before individuals were assigned to particular groups or categories using the program *Structure*. Finally, our third requirement was that the model populations must contain at minimum the final number of loci available for analysis for OUs within basins.

Input parameters for *MS*. *MS* is a Monte-Carlo program and therefore the same set of parameters can produce different results each run. After trial and error to produce the needed number of loci (see Table 3), we set mutation rate parameter to $4N_o\mu=20$, with N_o = diploid population size and μ = neutral mutation rate (Figure 2). Before and after population divergence, we set each ancestral or descendent population to have stable population growth (i.e., population size did not increase or decrease). To create the needed hierarchical relationships among the three populations we needed to establish two different divergence dates (dates corresponding to the temporal nodes A and B in Figure 2). We iteratively altered the divergence date input parameters until we produced three populations with F_{ST} values that approximated those among the wild, EWH and ESH steelhead OUs within basins in north Puget Sound.

Output from *MS* and subsequent population modeling. The output from *MS* was 500 haploid genotypes (essentially gametes) from each of the three populations (1500 total individuals). We then generated six "populations" representing source categories for *Structure* analyses: Pure Wild, pure EWH, pure ESH, hybrid: EWH-Wild, hybrid: ESH-Wild, hybrid: EWH-ESH. Here, we randomly paired, without replacement (i.e., monogamous mating) 200 gametes each from the Wild, EWH, and ESH populations generated by *MS*, to produce 100 diploid individuals each for the pure Wild, pure EWH, and pure ESH source categories (total of 300 individuals). Next, we again randomly paired, without replacement 200 gametes each from the Wild, EWH, and ESH populations generated by *MS*. These individuals represented the parental generation for our F1 hybrids. We randomly paired, without replacement each of 100 parents from the Wild, EWH, and ESH populations to produce 100 diploid individuals each for the hybrid: EWH-Wild, hybrid: ESH-Wild, hybrid EWH-ESH source categories, for a total of 300 hybrid individuals, and a grand total of 600 individuals, 100 each from the 6 source categories. We repeated this entire post-*MS* process 100 times, and chose from the 100 different datasets, the 10 datasets that best modeled the empirical F_{ST}s from the steelhead OUs (Table 4). Due to the similarity in F_{ST}s from the Green, Snohomish, Stillaguamish, and Skagit (including Samish) empirical datasets, we used one set of model populations for this Whidbey Basin (plus Green River) set, and another set of modeled populations for the Nooksack basin (Table 4). We used a combination of *Perl* and *Matlab* custom scripts to generate all modeled populations derived from the *MS* output.

Determining Structure thresholds and assignment error from modeled populations

Structure runs. The program *Structure* (Pritchard et al. 2000, Falush et al. 2003) is one of the most widely used programs for inferring population structure (see Gilbert et al. 2012 for summary of its use), and has also been used for detecting hybrid individuals, frequently between wild and domestic populations (e.g., Norén et al. 2005, Kidd et al. 2009, Sanz et al. 2009, Marie et al. 2011, Lamaze et al. 2012, Seamons et al. 2012, Harbicht et al. 2014). *Structure* makes use of each individual's multilocus genotype to infer population structure (e.g., hatchery versus wild), given an *a priori* assumed number of groups or populations (k). The program will probabilistically assign individuals to populations, or if the admixture option is used, will assign

a portion of an individual's genome to populations (Pritichard et al. 2000). We ran *Structure* on each of the ten population sets from the Whidbey and Nooksack models (Table 4) using admixture and maintaining credible regions, k = 2 – 4 populations, and three iterations for each population set and k. Initial runs were set at both 50,000 burn-in and 100,000 data collection chains (designated here, 50/100), and 5,000 burn-in and 50,000 data collection chains (5/50). Both sets of parameters provided the same results, so all subsequent runs were kept at the shorter 5/50 chains. Therefore, we ran *Structure* nine times (three ks with three iterations) for ten datasets for a total of 90 runs per model, times two models (Whidbey and Nooksack), for a total of 180 *Structure* runs.

Q-score thresholds. Our main target was k = 3 as our input was three populations (Wild, EWH, and ESH) and their hybrids. Since we selected the admixture option, *Structure* partitioned a portion of each modeled individual's genome into each of the three populations. These portions are represented by Q-scores and run from 0 (0% of the individual's genome) to 1 (100% of the individual's genome). We assigned each individual into one of seven assignment categories, representing the six source categories described above, and a No Call category where an individual could not be assigned with confidence (Figure 3). To establish the assignment regions in Figure 3, we varied the threshold value from 0.05 to 0.20 in 0.05 intervals, and used the following protocols: (1) identified an individual as pure Pop 1 (i.e., Wild) if the Q-scores for Pop 2 (EWH) and Pop 3 (ESH) were both less than the threshold value; (2) identified an individual as a hybrid between Pop1 and Pop 2 if the Q-score for Pop 3 was less than the threshold value and Q-scores for Pop 1 and Pop 2 were both greater than the threshold value, and (3) identified an individual as No Call (i.e., not assigned to an assignment category) if the Q-scores for all three populations were greater than or equal to the threshold value. Assignments to Pop 2 and Pop 3, and their hybrids, were assigned in like fashion. We selected the threshold value that minimized the overall assignment error rate (Table 5, Figure 3).

Once we established the appropriate threshold for individual assignments we needed to reconcile for each individual the assignments from the three iterations for each population set and k = 3. We established the final assignment for each individual using the following protocol. *Method 1*: For each individual we used the assignment that occurred in a minimum of 67% of the iterations (e.g., two of the three iterations). If no assignment occurred at a rate of 67%, but if one of the assignments was a hybrid (e.g., EWH-Wild) at any of the three iterations and the other assignments from the other iterations were one or both the pure categories included as part of the hybrid (e.g., EWH and/or Wild), we used the hybrid assignment. If neither of these conditions existed, we used No Call as the assignment. *Method 2*: We used the mean Wild, EWH, and ESH Q-scores across all iterations and assigned individuals to one of seven categories using the threshold method described above. If the assignments from Methods 1 and 2 above were the same, we used that assignment as the final assignment. If the assignment from Methods 1 and 2 were different, and one of the two assignments was a No Call, we used the non-No Call assignment as the final assignment. Otherwise, we examined the individual Q-scores for all iterations, and selected the assignment based on this overall evaluation of the Q-scores. The final assignment here was either that from Method 1 or 2.

Structure error rates. In a typical *Structure* analysis (i.e., not based on modeled populations), the source category (or population) is not known, and is typically what *Structure* is being used to estimate. Therefore, we evaluated the efficacy of the *Structure* runs using the assignment error

rate. We defined assignment error two ways. Source error rate is the frequency at which a source category is incorrectly assigned, either as a proportion of the total in that source category (i.e., proportion incorrectly assigned out of 1000 source individuals), or as a percentage of the total assigned (column assignment errors in Table 5). Assignment error rate is the proportion of individuals incorrectly assigned to a specific assignment category (row assignment errors in Table 5). An assignment category is proportionally over- or under-estimated by *Structure* when the total assigned to that category is greater than, or less than, respectively, the total that should be assigned (N = 1000) minus the unassigned (No Call) from that source category (Table 5).

Structure analysis with empirical data and assigning individuals to assignment categories

We used *Structure* to assign all steelhead individuals within each Wild OU (Table 1b) to one of the seven categories, as described above, using k = 2 through 4 or 5. For each Wild OU we included early winter hatchery (EWH) and early summer hatchery (ESH) samples (Table 2) in all *Structure* runs; however, EWH and ESH samples were not available in every basin. For all basins, except the Stillaguamish, a within-basin set of EWH samples were available and genotyped (Table 2). We used the Tokul Creek (Snohomish) samples as a surrogate EWH for *Structure* analyses with the Stillaguamish Wild OUs. ESH programs exist only in the Green, Snohomish, and Stillaguamish basins. We genotyped samples from Soos Creek (Green River) ESH program and the Reiter Ponds (Snohomish/Skykomish River) ESH program, and included these ESH samples along with the Wild and EWH samples in the Green and Snohomish *Structure* analyses. We also used the Reiter Ponds samples as the surrogate ESH OU in the Stillaguamish, Skagit, and Nooksack *Structure* analyses. The Samish River OU was run with the Skagit samples, and therefore included the Marblemount EWH and Reiter Ponds ESH samples (Table 2). In addition to the target Wild OU, EWH, and ESH samples in each *Structure* analysis, when possible, we also included in each *Structure* analysis another Wild (non-target) OU from the same basin as the target OU, resulting in a total of four OUs for most analyses, but aggregated into three source categories: Wild, EWH, and ESH).

<u>k = 3 analyses</u>. Although we evaluated k = 2 – 4 or 5, as with the modeled populations our explicit hypothesis was k = 3 (Wild, EWH, ESH source categories). We assigned individuals to k = 3 using threshold values and methods described above for the modeled populations, including both 5/50 and 50/100 chains. For each OU and k = 3, we repeated the analysis a minimum of 5 iterations and established the final assignment for each individual using Methods 1 and 2 described above for the modeled populations, except here for Method 1 the assignment rate to a category needed to be a minimum of 80% (e.g., four of the five iterations).

<u>k = 4 analyses</u>. The k = 3 *Structure* analyses were intended to generate three *Structure* groups that cleanly sorted most wild individuals into a Wild assignment category, EWH individuals into a EWH category, and ESH individuals in a ESH category. However, for some analyses individuals did not fall cleanly into these three categories. For example, EWH and ESH individuals would be assigned to a single "hatchery" category, and the wild individuals would be split into two different categories, making difficult the identification and interpretation of the assignment categories. Therefore, if the k = 3 analysis did not produce three clean Wild, EWH, and ESH categories, but the k = 4 analysis established EWH, ESH, Wild (either native winter or summer), and a Local category (i.e., none of the three standard categories), we used the k = 4

analysis, and assigned individuals completely to one of these four categories if Q-score for that category was ≥ 0.80. If no Q-score was ≥ 0.80, we assigned the individual as a hybrid between the categories with the largest and second largest Q-score if the second largest Q-score ≥ two times the third largest Q-score. Otherwise, the individual was assigned as No Call. We combined and reconciled the assignments based on the minimum of five iterations using methods described above for k = 3, except for Method 2 we used the k = 4 assignment procedure for the mean values, rather than the k = 3 procedure, based on the modeled populations.

Finally, for each OU, we totaled the number of individuals assigned to each assignment category, based either on the k = 3 or 4 analyses. We used these frequencies as the *Structure*-based summary assignments for the OU.

Assignment category nomenclature

We designated pure wild lineages as Basin Winter (i.e., occurs within more than one subbasin), or Local Winter or Summer (i.e., occurs within only one creek or river within the basin). There were no pure Basin Summer lineages. Pure hatchery lineages were designated as either Early Winter Hatchery (Chambers Creek – origin) or Early Summer Hatchery (Skamania – origin), and mixed lineages between pure wild and pure hatchery lineages were designed as Hybrid (see Table 6).

Identity of Local assignments

We initially assumed that individuals assigned as Local based on *Structure's* k=4 analyses were wild not hatchery lineage, because the individuals did not clustered with either EWH or ESH. To determine the historical relationship between Local and Wild categories within a basin, we used phylogenetic and multivariate analyses.

Data set: We included only "pure" categories (EWH, ESH, Wild, and Local) in the phylogenetic and multivariate analyses. To calculate allele frequencies of each of these pure groups, we attempted to strike a balance between being too restrictive by limiting the analyses only to those individuals defined by *Structure* as being pure (i.e., including no individuals in any of the hybrid categories), contrasted with being overly inclusive and introducing exogenous alleles into a pure group by conducting the analysis on the OUs themselves (i.e., introducing hatchery alleles into the Wild group or wild alleles into the hatchery groups). For these phylogenetic and multivariate analyses only, in each OU, we classified an individual as pure Wild (k = 3 or 4) or Local (k = 4 only) if it's Wild or Local Q-score was equal to or greater than 0.55 (Figure 4). A Q-score ≥ 0.55 would encompass 90% of the genetic variance of individuals whose true source was Wild (but may have been assigned to a hybrid category), 46% of the genetic variance of individuals whose true source was EWH-Wild (but may have been assigned as pure Wild), and 31%, 3%, 2%, and 0% of the genetic variance of individuals whose true source were ESH-Wild, EHW, EHW-ESW, and ESW, respectively.

Phylogenetic. We used a *Matlab* (MathWorks 2012) custom script to convert population allele frequencies to pairwise Nei genetic distances (Nei 1972), and then used *Mega* (Tamura et al. 2013) to construct a neighbor-joining tree from the distances. We used *Phylip* (Felsenstein 2005) to bootstrap the loci to construct 1000 new datasets, Nei distances, and trees and to

provide the frequency of any node that appeared in at least 1 of the 1000 trees. We determined the support for a node within the "true" tree constructed from the observed allele frequencies using the frequency of that node among the trees using bootstrap analysis. The historical relationships for each Local, Wild, EWH, and ESH category were assessed based on its phylogenetic relationships and the bootstrap support for that relationship.

Multivariate. We conducted a principal component analysis (PCA) to examine the genetic similarity among the "pure" EWH, ESH, Wild, and Local categories. The PCA was based on individual allele frequencies, and to assess relative genetic similarities we calculated pairwise Mahalanobis distances between the centroids for each Local group and the centroids for each EWH, ESH, and Wild group (regardless of basin). All analyses were conducted in *Matlab* (MathWorks 2012) using custom scripts.

Adjusting structure assignments using known assignment errors from model

The program *Structure* is known to make incorrect assignments under a variety of conditions (e.g., Vähä and Primmer 2006, Anderson and Durham 2008, Kalinowski 2011, Seamons et al. 2012). From our analyses of the modeled populations we know that *Structure* assigned modeled Puget Sound steelhead individuals to the incorrect category (i.e., source and assigned categories differ), with the assignment error rate dependent on the source and the assigned categories. Therefore, to correct for potential misassignments of empirical individuals, assessed across the entire OU, we needed to adjust the frequencies in the *Structure*-based summary assignments for each OU. In other words, the output from *Structure* provides assigned categories. We needed to convert these assigned categories to source categories, and we used the known assignment error rates from the model assignments to help with this conversion. We used a likelihood approach (Warheit and Knapp, in prep) to convert the *Structure*-based summary assignment to source assignments. This approach involved 15 steps, detailed below and summarized in Figure 5.

Step 1. Aggregated each of the assigned categories from the *Structure* analyses of the empirical data into one of the six model-based categories (Figure 3), as designated in Table 6. Individuals assigned to the No Call category were ignored.

Step 2. Calculated relative frequencies for each of the six categories. These relative frequencies can be used to estimate hatchery-wild introgression from the *Structure*-based estimates without correcting for assignment error.

Step 3. Based on relative frequencies from Step 2, expanded sample to a larger size, here we used 1000.

Step 4. Imported assignment error matrix from the modeled populations (Table 5). In Figure 5, the Simulated Error Matrix under Step 4 is the upper part from Table 5.

Step 5a. The relative frequencies for each of the six categories (Step 2) were adjusted separately with each category, in turn, being the designated target category. In Figure 5, we used as an example the Hybrid: EWH-Wild category as the target assigned category. The expanded count for the target category (EWH-Wild) is 208 (Step 3). This target category in the Simulated Error Matrix (Step 4) has 1162 individuals assigned to it. The actual source categories for these 1162 are the counts along the EWH-Wild row in Step 4 and in Table 5. We randomly selected, with replacement, 208 individuals from the 1162 modeled individuals assigned to the EWH-Wild category (designated here "Target Category").

Step5b. Repeated the process from Step 5a for the other assigned categories (designated here "Other Categories"). For example, randomly select with replacement, 680 of from the 931 modeled individuals assigned to the Wild category. Randomly selected individuals from this step were compiled together (as Other Category) but kept separate from the randomly selected individuals from Step 5a (Target Category).

Step 5 summary. The result from Step 5 was a new dataset of 1000 randomly selected modeled individuals from known source populations.

Step 6. The Simulated Error Matrix was based on *Structure* analyses where there were equal numbers of individuals from each of the six source categories (see Figures 1 and 2). These six categories are most-likely not equally represented in each of the OUs. In fact, it is these relative proportions that we are attempting to estimate. To simulate different relative proportions between the target and other categories, we constructed a series of new datasets, composed of the original sample size from Step 1 (for this example, N = 72), by randomly selecting specific number of Target and Other category individuals from the 1000 individuals compiled at Step 5. For example, a dataset that simulates 0% Target individuals and 100% Other individuals is composed of 72 randomly selected individuals from the N = 792 Other Category. A dataset that simulates 60% Target individuals and 40% Other individuals is composed of 43 randomly selected individuals from the N = 208 Target category and 29 randomly selected individuals from the N = 792 Other Category. This process was repeated for 0% to 100% Target Category, at 10% intervals, for a total of 11 new datasets, each composed of 72 individuals of know source category.

Step 7. For each of these 11 new datasets, we counted the number of individuals whose source category is the same as the Target assigned category (for the example, EWH-Wild). We then converted these counts to relative frequencies.

Step 8. Repeated Steps 6 and 7 multiple times. For this analysis we repeated this process 10,000 times to produce a 10,000 x 11 matrix composed of relative frequencies of individuals whose source is the Target category.

Step 9. To quantify the relationship between the assigned Target category proportion (0% - 100% at 10% intervals) and the source Target category proportion (from Step 8), we conducted a weighted least squares analysis. We used a weighted least squares approach because we assumed that the regression errors were correlated with each other and with the assigned Target category proportions, and had unequal variances. From this analysis, we recorded the intercept (β_o) and slope (β_1) of the regression line.

Step 10. To estimate the source Target category relative frequency from the assigned Target category relative frequency, we used the likelihood function for the normal regression,

$$\mathcal{L}\left(\beta_0, \beta_1, \sigma^2 \mid Y, X\right) = \frac{1}{\sqrt{(2\pi\sigma_i^{\,2})}} e^{-\frac{1}{2\sigma^2}(Y-(\beta_0+\beta_1 X_i))^2}$$

Equation 1

with,

Y = empirical assigned Target proportion (from Step 2; here 0.21),

X_i = fitted proportions (regression line in Step 9) from i= 0 to 1.0 in 0.001 intervals, with i being the estimated source Target proportions,

β_o and β_1 from weighted least squares (Step 9),

σ^2 from the least squares regression of variance (the "weights" from Step 9) against the assigned Target category proportions, and calculated as: $\sigma^2 = s\beta_o + (s\beta_1 * X_i)$, with $s\beta_o$ and $s\beta_1$ from regression, and X_i as above (See Figure 5).

Step 11. Used the function from Step 10 to calculate likelihoods for each X_i. The estimated source Target proportion, given an assigned Target proportion from *Structure* is the proportion where the likelihood is maximized (for this example, maximum likelihood is 8.85, with an estimated source Target proportion = 0.16, "correcting" the assigned proportion of 0.21 (from Step 2).

Step 12. Calculated confidence intervals (here, 90% CI) for the point estimate (i.e., the proportion where the likelihood is maximized) using the log-likelihood ratio test. In other words, we determined what likelihood values were not significantly different from the maximum likelihood value. We defined alpha as 0.10, with the critical value approximated using chi-square with 1 degree of freedom. The confidence interval was defined as the range of likelihoods that were not significantly different from the maximum, given alpha. In Figure 5, the 90% CI range was defined as those likelihoods that fell below the critical value, and the end points of the range were the smallest and largest likelihoods within that range.

Step 13. Adjusted the *Structure* assigned Target proportion, in this example, from 0.21 to a source Target proportion of 0.16, with a 90% CI = 0.00 to 0.39.

Step 14. Not shown in Figure 5. Since we adjusted or corrected each assigned category separately (i.e., each category is separately considered as Target, see Step 5a), the new source proportions, across all categories, for an OU may not sum to 1.00. Therefore, we divided each source proportion by the sum of all source proportions for that OU to give the final proportions for each category for the OU.

Step 15. Not shown in Figure 5. Finally, we reversed the aggregation of the assigned categories (Step 1), by proportionally de-aggregating each of the six modeled-based categories to produce the relative frequencies of source categories for each of the OUs.

Converting OU into NOAA PSSTRT Demographically Independent Populations (DIPs)

We provided the crosswalk between OUs and their respective DIPs in Figure 1b, and aggregated OUs into DIPs by adding together the frequencies in each *Structure* assigned category across all OUs within each DIP. For DIPs where spawning distributions are not evenly distributed among the contributing OUs, we weighted the OU's frequency for each category by the OU's estimated spawning proportion within the DIPs (Table 7). For each category and OU we multiplied the frequency by the appropriate weight (Table 7). Then for each category we added these products across all OUs contributing to that DIP. Because we applied a weight to each frequency the sum of the weighted products for each category was less than it would be if we had not applied the weights. To adjust the category sums so that the sample size for the entire DIP would be equal the sum of the sample sizes of the contributing OUs, we multiplied each category sum by the unadjusted sum of the contributing OUs across all categories, and then divided that product by the adjusted sum of the contributing OUs across all categories. We aggregated OUs into DIPs

first using only those OUs that were composed of adult fish, and then second, using all OUs, regardless of the life stage of the individuals.

As with the OUs, we needed to adjust or correct the frequencies in these newly constructed *Structure*-based summary assignments for each DIP. To do this we repeated Steps 1 – 15, described above for the OUs, for each adult- and all life stage- composed DIP.

Final summary statistics: effective pHOS and introgression

We provide two statistics that summarize the genetic contribution of hatchery lineage fish to each OU and DIP

Effective pHOS: The standard demographic definition of pHOS is the number of hatchery-origin (adipose fin clipped) fish spawning in the river divided by the total number of spawning fish. None of our wild or local fish were of hatchery-origin (i.e., all fish are adipose fin present); therefore, based solely on our samples demographic pHOS would be zero. However, natural-origin hatchery-ancestry fish (either pure EWH or ESH, or hybrids involving EWH and ESH) represent hatchery-origin (adipose fin clipped) fish that have spawned in the river in a previous generation. We estimated this genetic-based "ancestral" pHOS (hereafter effective pHOS) by assuming that the individuals assigned to a pure EWH and ESH category were products of two hatchery ancestors, while individuals assigned to a hybrid category that included EWH or ESH, and Wild (or Local) were products of one hatchery ancestor. To calculate effective pHOS we treated the number of hatchery ancestors as a weight (0, for no hatchery ancestors, 1, and 2, for 1 and 2 hatchery ancestors, respectively) and multiplied each category proportion for each OU or DIP by its appropriate weight (see Table 6). We summed the weighted proportions for each OU or DIP and divided that total by two (every individual has two parents). We calculated a separate effective pHOS for EWH and ESH, and interpreted these effective pHOS measures as the proportionate early winter or early summer hatchery ancestry of individuals from that OU or DIP

Introgression. The only avenue for a hatchery fish to directly genetically affect a wild population is to hybridize with individuals from that population. Therefore, by definition, to measure introgression within a population we are concerned only with hybrid fish, and introgression is a measure of the potential for direct hatchery effects on the allele frequencies of a wild population. To calculate introgression for either an OU or DIP, we added together the relative proportion of the hybrid categories that included Wild or Local (with some exceptions; see Results) and EWH or ESH categories (see Table 6 for which categories are included in the introgression calculation).

When there are no assigned pure EWH or ESH effective pHOS will be half the introgression. For example, if an OU was composed of 80% Wild winter and 20% hybrid: EWH-Wild, the effective pHOS would be 0.20/2 = 0.10 or 10%, and the introgression will be 20%. That is, 20% of the OU has been introgressed, but since these introgressed fish have only one parental hatchery lineage, the effective pHOS is 10%

Miscellaneous statistical analyses

Unless otherwise indicated, we conducted all statistical analyses in *Matlab* (MathWorks 2012) using custom scripts.

Results and Discussion

Samples and Loci

We genotyped 1728 Wild (natural-origin) fish from 47 collections and 33 OUs (Table 1), and 466 hatchery-origin fish from eight collections and six hatchery programs (Table 2). Of the 1728 Wild fish, we removed 34, 47, 328 individuals due to incomplete genotypes, presence of cutthroat alleles, and half-sib or closer relatedness, respectively. As expected, we found more related individuals among the juvenile collections (N = 251), than among the other life stages. The 251 individuals represented 34% of all juvenile samples genotyped, compared with only 8% from the adult samples being from related individuals. Of the 187 individuals removed from the hatchery collections, all but 4 individuals were removed due to half-sib or closer relatedness. There was a range of relatedness among the hatchery collections, from 60% of the individuals removed from the Kendall Creek – early winter collection to only 10% from the Reiter Pond early summer collection (Table 2). Across all Wild and hatchery OUs, we removed 27% of the individuals and used N = 1597 for statistical analyses.

Three of the 189 steelhead-specific loci were removed from all samples because fewer than 80% of all individuals had a usable genotype (N = 2) or the locus was monomorphic (N = 1). Of the remaining 186 loci, between 178 (Nooksack), 183 (Snohomish), and 184 (Stillaguamish, Skagit, Samish) loci were used for basin-specific analyses (Table 3).

Modeled populations

The "Before" and "After" F_{ST}s were similar to each other for each of the pairwise comparisons for the Nooksack, Whidbey Basin, and Green River OUs, except for the Green River EWH versus ESH comparisons (Table 4). The Before F_{ST} for this latter comparison was biased due to the inclusion of ESH fish in the Soos Creek – early winter collection (Figure 13a), which were removed for the After F_{ST}. The F_{ST}s among the Green, Snohomish, Stillaguamish, and Skagit operational groups were also similar to each other, and were closely matched by the F_{ST}s associated with each of the Whidbey Basin modeled populations, indicating that at least in terms of genetic differentiation, the modeled populations adequately represented the empirical populations from the Whidbey Basin and Green River (Table 4). The F_{ST}s for Nooksack modeled populations closely matched those from the Nooksack empirical populations, and, as with the Whidbey Basin model, the Nooksack modeled populations adequately represented the genetic differentiation of the empirical populations (Table 4). Overall, the Nooksack OUs were more differentiated from both EWH and ESH than the OUs from the other river basins. In addition, at least for the Whidbey Basin and Green River OUs, genetic differentiation between ESH-Wild was roughly the same as that for EWH-ESH, which is to be expected for neutral loci given the historical relationships among these three taxa, and the fact that Wild and EWH share a

more recent common ancestor with each other than either do with ESH (i.e., Wild and EWH are both native to Puget Sound steelhead, and ESH is native to lower Columbia River).

Model Structure assignment error

The overall assignment error rate was higher for the Whidbey Basin model (29%) compared with the Nooksack model (19%) (Table 5), which was expected, given the larger Nooksack F_{ST}s in all pairwise comparisons, compared with that from the other river basins (Table 4). For both models, the assignment category with the highest error rate was Hybrid:EWH-Wild (46% and 29%, Whidbey and Nooksack models, respectively), which was also expected, given the relatively low F_{ST} between EWH and Wild (Table 4). The largest source of error for the Hybrid:EWH-Wild category was the presence of large numbers of pure Wild and EWH fish in the Hybrid:EWH-Wild category (Table 5, Figure 6). The error rates for EWH (22% and 15%) and Wild (23% and 14%) categories were roughly half that for the Hybrid:EWH-Wild category, and more Wild (23% and 13%) and EWH (22% and 13%) fish incorrectly assigned as Hybrid:EWH-Wild than hybrid fish assigned to either the EWH (11% and 9%) or Wild (15% and 9%). The contribution of the other three categories (ESH and the hybrid categories involving ESH) to the total assignment to EWH, Wild, or Hybrid:EWH-Wild were either low and equal (8% for the Whidbey Model) or ≤ 5% (Nooksack Model). This means that the misassignments of EWH and Wild fish as Hybrid:EWH-Wild generally resulted in an overestimate of hybrid fish, and an underestimates of the pure Wild and EWH fish. The opposite seems to be true, but to a lesser degree for Hybrid:ESH-Wild assignments. That is, fewer Wild (3% and 5%) and ESH (2% and 5%) fish incorrectly assigned as Hybrid:ESH-Wild than hybrid fish assigned to either ESH (9% and 7%) or Wild (8% and 3%), but, these error rates are considerably lower than that associated with the EWH assignment categories (Table 5, Figure 6).

Structure assignment categories and the identity of the local assignments

With the exception of the Stillaguamish, the *Structure* analyses in all other river basins produced at least one pure Wild winter-run assignment category (Table 6). In the Stillaguamish at the time of our analyses, we had no sample collections composed exclusively of known Wild winter fish, and therefore, there was no Wild winter OU. In all but the Nooksack *Structure* analyses there were fish assigned to the pure ESH category, and, likewise, in all but the Stillaguamish *Structure* analyses there were fish assigned to the pure EWH category (Table 6). *Structure* assigned fish to pure Wild summer-run categories in the Stillaguamish (Canyon and Deer Creeks), Skagit (Finney Creek), and Nooksack (South Fork), but there were no wild-lineage summer-run fish assigned in the Snohomish basin, despite the presence of unmarked natural-spawning summer-run fish in the NF Skykomish and the SF Tolt rivers (see below). Overall, there were Local assignment categories (i.e., assignment categories limited to only a single watershed within the larger river basins) for either winter- or summer-run fish in the Snohomish (Pilchuck, NF Skykomish, and SF Tolt), Stillaguamish (Deer and Canyon), Skagit (Nookachamps and Finney), and Nooksack (SF Nooksack) (Table 6). Additionally, there was a Local assignment category in the NF Tolt; however, subsequent analyses (not shown in this report) indicated that the juvenile fish assigned to this local category were not steelhead, and were probably non-native rainbow trout. These six fish were removed from all subsequent analyses.

The local summer-run fish from the Snohomish system (all three NF Skykomish and the SF Tolt Local categories) were phylogenetically placed within the ESH group, indicating that they shared a more recent common ancestor with the lower Columbia River Skamania-lineage group than they did with Puget Sound summer- or winter-run fish (node 1, Figure 7). Furthermore, this phylogenetic relationship is robust, occurring in 86% of the 1000 bootstrap trees. By contrast, the local summer-run fish from the Stillaguamish (Deer and Canyon) and Skagit (Finney) cluster together (node 8) and are phylogenetically placed well-within the Puget Sound wild steelhead lineage (node 3, Figure 7), which is also well supported by 88% of the 1000 bootstrap trees. The Pilchuck Local winter-run group appropriately clusters with other parts of the Snohomish lineage (node 7), and the SF Nooksack Local summer group occurs well within the Nooksack lineage (node 9), which is supported by 99% of the 1000 bootstrap trees. The phylogenetic position of the Nookachamps Local winter-run group is ambiguous. The neighbor-joining tree from Nei distances showed the Nookachamps Local winter-run group nested within a poorly defined Whidbey basin plus Nooksack River group (node 6), while the bootstrap consensus tree places the Nookachamps Local winter-run group between nodes 2 and 3, that is, between the group consisting of all Puget Sound OUs and the group consisting of all wild Puget Sound OUs (Figure 7).

Overall, the phylogenetic analysis strongly supports the monophyly of the ESH (node 1), EWH (node 4), Puget Sound Wild (node 3), Green River (node 5), Stillaguamish (node 8, but limited to Deer and Canyon Creeks, and also includes Finney Creek form the Skagit basin), the Nooksack plus Samish Rivers (node not labeled), and the Nooksack River (node 9) lineages. Also supported, but at lower bootstrap support is the Snohomish River lineage. The phylogenetic analysis does not support the monophyly of the Skagit River group, indicating that the "lineage" is both polyphyletic (i.e., the separation of the Finney summer-run, and the Sauk and Suiattle group from the rest of the Skagit), and paraphyletic (Figure 7).

Our principal component analysis of these taxa was consistent with the phylogenetic analysis (Figure 8). That is, ESH, EWH, Green River, and Nooksack River groups formed distinct and tight clusters. The local summer-run fish from the Snohomish system, which were phylogenetically placed with ESH were intermediate between ESH, EWH, and the wild taxa. Mahalanobis distances place the NF Skykomish Local summer-run fish closer to ESH than the other groups, but SF Tolt Local summer-run appeared to be equally distant between ESH and EWH (Figure 9). Deer and Finney, and less so Canyon Local summer-run groups were separated from the Wild group, in the direction towards ESH, but their Mahalanobis distances were clearly shorter with the Wild group than either EWH or ESH groups (Figure 9). Likewise, the Pilchuck and Nookachamps Local winter-run groups, and the SF Nooksack Local summer group were somewhat separate from the remainder of the Snohomish, Skagit, and Nooksack groups, respectively, but their Mahalanobis distances placed these fish unmistakably with the Wild group (Figure 9).

Based on the phylogenetic and multivariate analyses, we considered the NF Skykomish and SF Tolt Local summer-run fish to be of non-local origin, and phylogenetically part of ESH – lower Columbia clade. However, it is possible that although the majority of the genome for each of these fish is of ESH origin, their overall genetic makeup may be a consolidation of alleles originating in ESH, EWH, and Wild groups, including a potentially extinct Wild summer group.

This may be particularly true for SF Tolt Local summer-run, where, based on the principal component analysis (Figure 8), the group appears equally distance between ESH and EWH. We consider all other local assignments to be Puget Sound wild groups.

Genetic composition of OUs and demographically independent populations

We present the *Structure* proportions for each basin's OUs and DIPs Figures 7 – 12. These results are shown graphically along with the category-specific likelihood-adjusted proportions (i.e., the results from Steps 1 – 13) in Figure 5. In most cases the *Structure* proportions are within the 90% CI of the adjusted proportions, but not always. The direction and amount of the adjustment (i.e., difference between the *Structure* and adjusted results) can be determined by examining the relationship of the *Structure* results (green vertical line) and adjusted results (black vertical line) within the subplots in Figures 7 – 12. Confidence in both the *Structure* proportions and the adjusted proportions can be assessed by the width of the 90% CI, with a broad interval connoting low confidence and a narrow interval high confidence. We show the final adjusted proportions for each basin's OUs and DIPs in Tables 7a – f. These final adjusted proportions are described in Steps 14 – 15 (Methods section) and are the category-specific likelihood-adjusted proportions, when they sum to one across all categories, within an OU or DIP.

Green River

Preliminary analyses of the Green River OUs suggested that EWH unit (Soos Creek – early winter [Chambers]) was a mix between the EWH and the ESH (Soos Creek – early summer [Skamania]) programs. As such, we included here the composition of both the Wild and hatchery OUs. We provide the *Structure* proportions and the category-specific likelihood-adjusted proportions, with their 90% CI, for the OUs in Figure 13a, and for the DIPs in Figure 13b. The final adjusted proportions for both the OUs and the DIPs are in Table 6a, and effective pHOS and introgression summary statistics are in Table 8a.

OU assignments: The four Green River Wild OUs showed two different genetic compositions. The 2004 and 2013 adult units were composed entirely Wild winter fish, with no genetic influence of either the early winter or summer programs (Table 6a, 8a). However, the 2007 and 2008 smolt units were composed of roughly 0.25 Hybrid:EWH-Basin Winter, 0.73 Basin Winter, and 0.01 EWH (GreenR07) and 0.02 ESH (GreenR08).

DIP assignments: When considering the four OUs together as a single DIP, 0.88 were Wild winter, 0.11 were Hybrid:EWH-Basin Winter, and 0.01 were ESH (Table 6a).

Effective pHOS and introgression: The two smolt OUs showed a total effective pHOS of 0.15, nearly all from the early winter program, and a total introgression of roughly 0.25. The total effective pHOS for Green River Winter-Run DIP was 0.06, 0.05 from the early winter and 0.01 from the early summer programs. Total introgression was 0.11 (Table 8a). As expected from the preliminary analyses the Soos Creek – early winter collection included a considerable number of pure early summer program fish (0.32), while the Soos Creek – early summer collection was composed nearly entirely (0.99) of summer program fish.

Snohomish River

We provide the *Structure* proportions and the category-specific likelihood-adjusted proportions, with their 90% CI, for the Snohomish River OUs in Figure 14a, and for the DIPs in Figure 14b. The final adjusted assignments for both the OUs and the DIPs are in Table 6b and effective pHOS and introgression summary statistics are in Table 8b. The Snohomish River OUs and DIPs were the genetically most complex of all the river basins and showed the highest overall levels of introgression, and the highest effective pHOS statistics, with, for the most-part, the effective pHOS being affected mostly by the early summer program (Reiter Ponds).

OU assignments: The NF Skykomish and SF Tolt Above OUs were dominated by ESH-linage fish, either in the form of pure ESH, as ESH-ancestry Local Summer, or Hybrid: ESH- ESH-ancestry Local Summer fish (Table 6b). The ESH program (Reiter Ponds) also affected the composition of winter-run OUs in the Snohomish, with either ESH or hybrids that include ESH occurring in NF Tolt, SF ToltBelow, and mainstem Skykomish. Except for SF ToltBelow, most of the fish in the winter-run OUs were assigned to the Basin Winter category, with proportions ranging 0.50 – 0.76. For SF ToltBelow, the assignment category with the highest proportion was indeed Basin Winter, but the proportion here was 0.42, indicating that the majority of the fish were assigned elsewhere. In the Pilchuck OU, all fish assigned to winter categories; split among Basin Winter, Pilchuck Local Winter, and Hybrid:Basin Winter – Pilchuck Local Winter. Although none of the fish in the winter-run OUs were assigned to pure EWH, a large proportion of the fish from NF and SF ToltBelow, and Snoqualmie OUs assigned to Hybrid: EWH – Basin Winter.

DIP assignments: The only DIP that was not redundant with one of the OUs was the all sample Snoqualmie River Winter-Run DIP, which aggregated winter-run samples from Snoqualmie, and SF ToltBelow, and NF Tolt OUs. Here, 0.71 assigned to the Basin Winter category, 0.28 to Hybrid:EWH - Basin Winter category, and 0.01 to the pure ESH category.

Effective pHOS and introgression: For all but the NFSkyJuv04 category, all summer-run categories had effective pHOS – early summer = 1.00, indicating that 100% of the summer-run natural spawners in the NF Skykomish and SF ToltAbove were hatchery (ESH) – derived fish, as indicated by these OUs. NFSkyJuv04 showed a total effective pHOS = 1.00, but this was split 0.19 early winter and 0.81 early summer. Total introgression among these summer-run OUs ranged 0.26 (NFSkyJuv04) to 0.64 (NFSkySumAd2013). With the exception of the Pilchuck, all the winter-run OUs had total effective pHOS greater than 0.10. For the NF ToltAbove and mainstem Skykomish winter OUs, the total effective pHOS was composed entirely of the effective pHOS from the early *summer* program; 0.21 and 0.24, respectively. The SF ToltBelow OU had a total effective pHOS = 0.40, split 0.13 early winter and 0.26 early summer. Although the effective pHOS – early summer was high (0.24), introgression for this OU was zero, as the effective pHOS was a function of a single individual assigned as a pure ESH. Furthermore, the 90% CI for this categorical assignment ranged 0.06 to 1.00 (Figure 14a), indicating to us that we should have little confidence in this assignment and its resulting effective pHOS. For the other winter-run OUs, introgression ranged from zero (Pilchuck) to 0.37 (NF ToltAbove). The total effective pHOS and introgression for the all sample Snoqualmie River Winter-Run DIP were 0.15 and 0.27, respectively, with the effective pHOS mostly from the early winter program (0.13).

Stillaguamish River

We provide the *Structure* proportions and the category-specific likelihood-adjusted proportions, with their 90% CI, for the Stillaguamish River OUs in Figure 15a, and for the DIPs in Figure 15b. The final adjusted assignments for both the OUs and the DIPs are in Table 6c and effective pHOS and introgression summary statistics are in Table 8c. For this project our Stillaguamish samples were limited to Deer and Canyon Creek summer-run and an aggregate smolt-trap collection (Tables 1 – 2). We had no known within-basin winter-run samples to include in any of the *Structure* analyses of Stillaguamish OUs. In addition, we had no EWH or ESH collections from the Whitehorse facility on the Stillaguamish, and used as surrogates the Tokul Creek (EWH) and Reiter Ponds (ESH) facilities on the Snohomish. As such, the following results may be biased by the lack of a known winter-run OU, and by the fact that the composition of our existing OUs and DIPs were based on the allelic diversity of EWH and ESH from another river basin. However, the close genetic similarities among hatcheries within both pure ESH and pure EWH (Figures 13 – 14) suggest that our use of surrogates should not greatly bias our results.

OU assignments: Canyon Creek and Deer Creek summer-run OUs are distinct and assigned to different Local Summer categories. For Canyon Creek summer-run OU all individuals assign to wild categories, split between Local Summer Canyon (0.69) and Hybrid:Local Summer Canyon – Local Summer Deer (0.31). Although the Deer Creek juvenile 1995 OU showed a 0.24 hatchery hybrid proportion (0.21 Hybrid:ESH – Local Summer Deer, and 0.03 Hybrid:EWH - ESH), the Deer Creek juvenile 2013 OU was composed entirely of wild fish, divided into Local Summer Deer (0.82), Local Summer Canyon (0.02), and Hybrid:Local Summer Canyon – Local Summer Deer (0.16). The Stillaguamish smolt trap OU was composed of 0.88 wild categories, either Local Summer Canyon, Local Summer Deer, or their hybrids. There were no pure EWH, but 0.09 pure ESH, and 0.04 ESH hybrids, divided between Local Summer Canyon, Local Summer Deer, and and EWH. Some of these assignments may be incorrect since there were no wildwinter OUs included in the analyses

DIP assignments: When juvenile and adults samples were taken together as the Deer Creek Summer-Run DIP, the Hybrid:ESH-Local Summer Deer proportion declined to 0.02, with the remainder of the individuals assigned as wild, partitioned as Local Summer Deer (0.85), Local Summer Canyon (0.01), and Hybrid:Local Summer Cayon – Local Summer Deer (0.11). .

Effective pHOS and introgression: Effective pHOS and introgression were highest in the Deer Creek juvenile 1995 (0.13 and 0.21, respectively) and the Deer Creek adult (0.11 and 0.23)OUs; however, the Deer Creek adult OU included only eight individuals and no definitive conclusions should be drawn from that low of a sample size. The all-samples Deer Creek Summer-Run DIP, which included 1995 and 2013 juvenile, and the adult samples, showed a 0.01 total effective pHOS and 0.02 introgression. Except for Deer Creek juvenile 1995's 0.01 effective pHOS – early winter, all effective pHOS values reflected the early summer program.

Skagit River

We provide the *Structure* proportions and the category-specific likelihood-adjusted proportions, with their 90% CI, for the Skagit River OUs in Figure 13a, and for the DIPs in Figure 13b. The final adjusted assignments for both the OUs and the DIPs are in Table 6d and effective pHOS and introgression summary statistics are in Table 8d. Unlike the Green, Snohomish, and Stillaguamish Rivers, the Skagit River does not have an ESH program, but we included in the

Structure analyses the Snohomish Reiter Pond samples to search for an ESH-lineage genetic signal.

OU assignments – Finney Creek: We detected two Skagit River Local categories: Nookachamps Local Winter and Finney Creek Local Summer. Based on our phylogenetic analysis the fish assigned to Finney Creek Local Summer are more closely related to Local Summer Deer and Local Summer Canyon fish (Stillaguamish) than to any of the winter categories in the Skagit River, including the Finney Creek winter OU. Nonetheless, most of the individuals from Finney Creek summer (0.57) assigned to the Hybrid: Basin Winter – Finney Creek Local Summer category, as did one-third (0.32) of the Finney Creek winter OU individuals. Furthermore, 0.13 of the Finney Creek winter OU individuals assigned to the Finney Creek Local Summer category, 0.02 to pure EWH, but the majority (0.53) assigned to the Basin Winter category. The only signal for the occurrence of ESH-lineage fish in the Skagit River was in the Finney Creek summer OU, where 0.12 assigned to the Hybrid:ESH – Finney Creek Local Summer category. Taking the Finney Creek winter and summer OUs together, 0.44 of the fish in the Finney Creek basin assigned to the Hybrid:Basin Winter – Finney Creek Local Summer category, 0.22 to Finney Creek Local Summer category, 0.26 to the Basin Winter category, and 0.06 to the Hybrid:ESH – Finney Creek Local Summer category. The Hybrid:Basin Winter – Finney Finney Creek Local Summer category does not appear to be a "false category" that existed because we lacked the power to differentiate the winter and the local summer categories. Although we did not explicitly test the power to differentiate Finney Creek Local Summer from the Basin Winter individuals, the two categories are genetically distinct. First, *Structure* analyses clearly separated the two categories. Second, the average F_{ST} between the Finney Creek summer OU and the winter OUs in the Skagit River basin was 0.023, including a $F_{ST} = 0.028$ with the Finney Creek winter OU. Third, the Finney Creek Local Summer category did not cluster with any of the Skagit River winter categories in the principal component analysis (Figure 7).

OU assignments – Winter: Except for the Cascade River winter OU, all other winter OUs appeared to be completely or nearly completely composed of Basin Winter fish. In the Cascade River winter OU, 0.17 assigned as Hybrid:EWH – Basin Winter and 0.83 assigned to pure Basin Winter. Marblemount Hatchery, the source of EWH fish in the Skagit River is located near the mouth of the Cascade River, so the higher incidence of Hybrid:EWH – Basin Winter fish in the Cascade River was not surprising. In the upper Skagit adult OU, 0.01 assigned to pure EWH category. Finally, the Nookachamps OU was composed equally of Nookachamps Local Winter, and Hybrid:Basin Winter – Nookachamps Local Winter (0.49 each). The remainder of the OU assigned to the Basin Winter category.

DIP assignments: The Mainstem Skagit R Summer- and Winter-Run DIP was composed of 0.91 Basin Winter, 0.02 Finney Creek Local Summer, 0.05 Hybrid:Basin Winter – Finney Creek Local Summer, and 0.01 pure EWH. By contrast the Sauk R Summer- and Winter-Run DIP was composed entirely of Basin Winter fish.

Effective pHOS and introgression: The two larger effective pHOS values in the Skagit River belong to the Cascade River winter (0.08, all to the early winter program) and the Finney Creek summer (0.06 all to the early summer program). These two OUs were also the only units to show introgression (0.17 and 0.12, respectively). The only other signal of potential hatchery effects on the wild OUs in the Skagit River is an effective pHOS – early winter = 0.01 for the upper Skagit adult OU.

Samish River

We provide the *Structure* proportions and the category-specific likelihood-adjusted proportions, with their 90% CI, for the Samish River OU/DIP in Figure 14. The final adjusted assignments for both the OU/DIP are in Table 6e and effective pHOS and introgression summary statistics are in Table 8e. There is only one OU in the Samish River, which is equivalent to the DIP. The Samish River does not have a steelhead hatchery program but did receive hatchery plants from the Whatcom Hatchery in 2005, 2006, and 2007. We analyzed the Samish samples along with the Skagit River OUs, and they assigned 0.83 to Basin Wild and 0.17 Hybrid:EWH – Basin Wild. The effective pHOS – early winter was 0.08, and introgression was 0.17.

Nooksack River

We provide the *Structure* proportions and the category-specific likelihood-adjusted proportions, with their 90% CI, for the Nooksack River OUs in Figure 15a, and for the DIPs in Figure 15b. The final adjusted assignments for both the OUs and the DIPs are in Table 6f and effective pHOS and introgression summary statistics are in Table 8f. As with the Skagit River, the Nooksack River does not have an ESH program, but we included in the *Structure* analyses the Snohomish Reiter Pond samples to search for an ESH-lineage genetic signal.

OU assignments: We detected a local summer category in the Nooksack River; however, here, unlike the Finney Creek summer OU where most of the individuals assigned to a hybrid category, the SF Nooksack summer OU was 0.90 pure Local Summer, 0.02 Basin Winter, and 0.09 Hybrid:Basin Winter – Local Summer. There was no ESH signal in the SF Nooksack summer OU. The SF Nooksack winter OU assigned 0.97 to wild categories, within 0.70 to the Basin Winter, 0.03 pure Local Summer, 0.24 Hybrid:Basin Winter – Local Summer, and 0.03 EWH. Other OUs with pure EWH assignments were Mainstem early adults and NF Nooksack adults, and 0.08 of Mainstem early adults also assigned to Hybrid:EWH – Basin Winter category. We also detected an ESH signal in the Hybrid:ESH – Basin Winter category assignment in the Mainstem early adults (0.03) and NF Nooksack juveniles (0.10), and in the Hybrid:EWH – ESH category assignment in the Mainstem early adults (0.01).

DIP assignments: All Nooksack winter OUs were lumped into a single DIP, the Nooksack R Winter-Run DIP. Although this DIP is composed of mostly Basin Winter fish (0.91 and 0.89, adults only and all samples, respectively), individuals were also assigned to all other categories, except the Hybrid:EWH – Basin Winter category.

Effective pHOS and introgression: Total effective pHOS among the Nooksack OUs ranged from zero (SF Nooksack summer) to 0.09 (Mainstem early adults), with two OUs with effective pHOS – early winter only, one OU with effective pHOS – early summer only, and Mainstem early adults with 0.07 effective pHOS – early winter, and 0.02 effective pHOS – early summer. Total introgression was limited to Mainstem early adults (0.11) and NF Nooksack juveniles (0.10) only. These effective pHOS levels were also reflected in the Nooksack R Winter-Run DIP – all samples, with 0.03 effective pHOS – early winter and 0.02 effective pHOS – early summer. Total introgression in the DIP was 0.02.

Proximate causes for inter-basin differences in introgression

The amount of hatchery – wild introgression varied among the six river basins included in our analyses, from a high level in the Snohomish basin to a relatively low level in the Skagit and

Nooksack basins. As a *preliminary* examination of some of the potential causes of these differences, we considered total number of off-station hatchery releases and wild escapement as independent variables, and introgression as the dependent variable. We restricted the release data to off-station releases by assuming that these fish were the most-likely to not return to the hatchery (i.e., on-station) and therefore, spawn naturally with each other and with wild fish. Based on the availability of release and escapement data, we limited our analyses here to Green River winter-run, Snohomish River winter- and summer-run, considered separately, Skagit River winter-run, and Nooksack River total-run. For the off-station hatchery releases we used data from years 2000 – 2010, which correspond to return years 2002 – 2012. Therefore, for the wild escapement we also used return years 2002 – 2012. For each basin and run, we used only those years where data existed for both the hatchery releases and wild escapements, and for each year in common we created a third independent variable: ratio between average number of off-station releases and wild escapement. We represented the three independent variables for each basin and run using mean scores across all available years. The number of years included in the analysis varied among the basins, with Nooksack being the lowest (N = 4), Green River the highest (N = 11), and Snohomish and Skagit rivers with N = 10. For this analysis we calculated separate EWH and ESH introgression for each operational unit and DIP, by using only those hybrid categories that include EWH and ESH, respectively. We calculated two summary introgression scores for each basin and run, using the mean value across all (1) operational units and (2) DIPs – all samples.

There is no apparent relationship between introgression and the average number of off-station releases within a basin (Figure 16). However, there is a compelling relationship between introgression and the ratio between average number of off-station releases and wild escapement; the more off-station releases per wild escapement, the greater the amount of introgression (Figure 16). There is also a weak negative relationship between introgression and wild escapement (r^2 = 0.19 and 0.31 for the operational units and DIPs – all samples, respectively). This analysis suggests that introgression is not simply the result of having a large number of hatchery fish released off-station, but is also a function of the number of wild fish on the spawning grounds.

Conclusions

Based on samples used in the analyses here, segregated hatchery programs in the Green, Snohomish, Stillaguamish, Skagit, and Nooksack River basins have genetically affected the wild populations in these basins through introgressive hybridization. However, the amount of introgression varied within and among basins from small to extensive, with this variation possibly being a function, in part, of the number of off-station hatchery releases and wild escapement in each basin.

Acknowledgements

We thank Sarah Bell (WDFW-MGL) for all laboratory work associated with this project, and Todd Kassler (WDFW-MGL) for overall project management and sample curation. We also thank all members of the WDFW-MGL, in particular Todd Seamons, for discussions that helped focused this project. Todd Seamons also read several version of this report and provided helpful

comments and edits. Shannon Knapp (formerly WDFW, now New Mexico State University) designed the first version of the likelihood method to adjust structure assignments. This technical report is part of a larger effort to develop Hatchery Genetic Management Plans (HGMPs) for Puget Sound steelhead hatchery programs. The development of many of the concepts in this report was part of a coordinated effort among Jim Scott, Annette Hoffmann, Kelly Cunningham, Brian Missildine, and Beata Dymowska. Finally, this project was funded in part by Seattle City Light, and the Washington General and Wildlife Funds.

Literature Cited

Abadia-Cardoso, A., AJ Clemento, JC Garza (2011). Discovery and characterization of single-nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. *Molecular Ecology Resources* 11:31–49.

Aguilar A, JC Garza (2008). Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). *Molecular Ecology Resources* 8:659-662.

Anderson, EC, and KK Dunham (2008). The influence of family groups on inferences made with the program Structure. *Molecular Ecology Resources* 8: 1219-1229.

Araki, H, B Cooper, and MS Blouin (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science* 318: 100-103.

Blouin MS (2003). DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology & Evolution* 18:503-511

Brunelli JP, GH Thorgaard, RF Leary, JL Dunnigan (2008). Single-nucleotide polymorphisms associated with allozyme differences between inland and coastal rainbow trout. *Transactions of the American Fisheries Society* 137:1292-1298.

Campbell, NR, SR Narum (2009). Identification and characterization ofheat shock response–related single-nucleotide polymorphisms in *O. mykiss* and *O. tshawytscha*. *Molecular Ecology Resources* 9:1460–1559.

Christie, MR, ML Marine, RA French, RS Waples, and MS Blouin (2012). Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. *Heredity* 109: 254-260.

Cockerham, CC (1969). Variance of gene frequencies. *Evolution* 23:72-84

Crow, JF, and M Kimura (1970). *An Introduction to Population Genetics Theory*. New York, Harper and Row.

Falush D, M Stephens, JK Pritchard (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587

Felsenstein, J (2005). *PHYLIP (Phylogeny Inference Package) version 3.6*. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Finger, A, MR Stephens, NW Clipperton, B May (2009). Six diagnostic single nucleotide polymorphism markers for detecting introgression between cutthroat and rainbow trouts. *Molecular Ecology Resources* 9: 759-763.

Frankham, R, JD Ballou, and DA Briscoe (2002). *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.

Gilbert, KJ, RL Andrew, DG Bock, MT Franklin, NC Kane, J-S Moore, BT Moyers, S Renaut, DJ Rennison, T Veen, and TH Vines (2012). Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology* 21: 4925-4930.

Harbicht AB, M Alshamlih, CC Wilson, DJ Fraser (2014). Anthropogenic and habitat correlates of hybridization between hatchery and wild brook trout. *Canadian Journal of Fisheries and Aquatic Sciences* 71: 1–10 dx.doi.org/10.1139/cjfas-2013-0460

Hansen, MHH, S Young, HBH Jorgensen, C Pascal, M Henryon, J Seeb (2011). Assembling a dual purpose TaqMan-based panel of single-nucleotide polymorphism markers in rainbow trout and steelhead (*Oncorhynchus mykiss*) for association mapping and population genetics analysis. *Molecular Ecology Resources* 11:67–70.

Hess, MA, CD Rabe, JL Vogel, JJ Stephenson, DD Nelson, and SR Narum (2012). Supportive breeding boosts natural population abundance with minimal negative impacts on fitness of a wild population of Chinook salmon. *Molecular Ecology* 21(21): 5236-5250.

Hilborn, R (1992). Hatcheries and the future of salmon in the Northwest. *Fisheries* 17(1): 5-8.

HSRG (2004). *Hatchery Reform: Principles and Recommendations of the HSRG*. Long Live the Kings, Seattle. (available from www.hatcheryreform.org).

Hudson, RR (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 18(2): 337-338.

Kalinowski ST, AP Wagner, ML Taper (2006). ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes* 6:576-579.

Kalinowski ST (2011). The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity* 106:625-632.

Kidd, AG, J Bowman, D Lesbarrères, and AI Schulte-Hostedde (2009). Hybridization between escaped domestic and wild American mink (*Neovison vison*). *Molecular Ecology* 18: 1175-1186.

Kostow, K (2009). Factors that contribute to the ecological risks of salmon and steelhead hatchery programs and some mitigating strategies. *Reviews in Fish Biology and Fisheries* 19:9-31.

Laikre, L, MK Schwartz, RS Waples, N Ryman, and The GeM Working Group (2010). Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. *Trends in Ecology and Evolution* 25: 520-529.

Lamaze, FC, C Sauvage, A Marie, D Garant, and L Bernatchez (2012). Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (*Salvelinus fontinalis*). *Molecular Ecology* 21: 2877-2895.

Limborg, MT, SM Blankenship, SF Young, FM Utter, LW Seeb, MHH Hansen, JE Seeb (2011). Signatures of natural selection among lineages and habitats in *Oncorhynchus mykiss*. *Ecology and Evolution*: doi: 10.1002/ece3.59.

MathWorks Inc. (2012). *Matlab* R2012B (8.0.0.783). Natick, MA, 2000.

Marie, AD, L Bernatchez, and D Garant (2011). Empirical assessment of software efficiency and accuracy to detect introgression under variable stocking scenarios in brook charr (*Salvelinus fontinalis). Conservation Genetics* 12: 1215-1227.

McGlauflin M, MJ Smith, JT Wang, SF Young, N Chen, YC Lee, C Pascal, LW Seeb, J Stevens, JE Seeb (2010). High resolution melting analysis for the discovery of novel SNPs in rainbow (*Oncorhynchus mykiss*) and cutthroat trout (*Oncorhynchus clarki*) for species identification . *Transactions of the American Fisheries Society* 139(3):676-684.

Mobrand, LE, J Barr, L Blankenship, DE Campton, TTP Evelyn, TA Flagg, CVW Mahnken, LW Seeb, PR Seidel, and WW Smoker (2005). Hatchery reform in Washington State. *Fisheries* 30(6): 11-23.

Naish, KA, JE Taylor III, PS Levin, TP Quinn, JR Winton, D Huppert, and R Hilborn (2008). An evaluation of the effects of conservation and fishery enhancement hatcheries on wild populations of salmon. *Advances in Marine Biology* 53(61-194).

Nei M. (1972). Genetic distance between populations. *American Naturalist* 106:283-292.

Norén, K, L Dalén, K Kvaløy, and A Angerbjörn (2005). Detection of farm fox and hybrid genotypes among wild arctic foxes in Scandinavia. *Conservation Genetics* 6: 885-894.

Pritchard, JK, M Stephens, P Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Sánchez, CC, TPL Smith, RT Wiedmann, RL Vallejo, M Salem, J Yao, CE Rexroad (2009). Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. *BMC Genomics* 10:559.

Sanz, N, RM Araguas, R Fernández, M Vera, and J-L García-Marín (2009). Efficiency of markers and methods for detecting hybrids and introgression in stocked populations. *Conservation Genetics* 10: 225-236.

Seamons, TR, L Hauser, KA Naish, and TP Quinn (2012). Can interbreeding of wild and artificially propagated animals be prevented by using broodstock selection for divergent life history. *Evolutionary Applications* 5(7): 705-719.

Sprowles, AE, MR Stephens, NW Clipperton, BP May (2006). Fishing for SNPs: a targeted locus approach for Single Nucleotide Polymorphism discovery in rainbow trout. *Transactions of the American Fisheries Society* 135:1698-1721.

Stephens MR, NW Clipperton, B May (2009). Subspecies-informative SNP assays for evaluating introgression between native golden trout and introduced rainbow trout. *Molecular Ecology Resources* 9, 339-343.

Tamura K, G Stecher, D Peterson, A Filipski, and S Kumar (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.

Thériault, V, GR Moyer, LS Jackson, MS Blouin, and MA Banks (2011). Reduced reproductive success of hatchery coho salmon in the wild: insights into most likely mechanisms. *Molecular Ecology* 20(9): 1860-1869.

Vähä, J-P, and CR Primmer (2006). Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different number of loci. *Molecular Ecology* 15: 63-72.

Waples, RS (1991). Genetic interactions between hatchery and wild salmonids - lessons from the Pacific Northwest. *Canadian Journal of Fisheries and Aquatic Sciences* 48: 124-133.

Zhivotovsky, LA, LK Fedorova, GA Rubtsova, MV Shitova, TA Rakitskaya, VD Prokhorovskaya, BP Smirnov, AM Kaev, VM Chupakhin, VG Samarsky, VP Pogodin, SI Borzov, and KI Afanasiev (2012). Rapid expansion of an enhanced stock of chum salmon and its impacts on wild population components. *Environmental Biology of Fishes* 94: 249-258.

Table 1a. Geographic and temporal scope, and biological and management descriptors of wild (natural-origin or unmarked) steelhead collections used in this study.

Basin	Subbasin	Collection Code	Collection Year	Life Stage	Collection Dates	Origin	Presumed Run Timing
Nooksack	Mainstem	11NW	2011	adult	Dec 2010 - Jan 2011	wild	winter
Nooksack	Mainstem	12MP	2012	adult	Dec 2011 - Jan 2012	wild	winter
Nooksack	Mainstem	13GC	2013	adult	Dec 2012 - Jan 2013	wild	winter
Nooksack	Northfork	12MQ	2012	adult	Feb - April	wild	winter
Nooksack	Northfork	09MN	2009	juvenile	Fall	wild	winter
Nooksack	Northfork	10PY	2010	juvenile	unknown	wild	winter
Nooksack	Southfork	12CF	2012	adult	Feb - March	wild	winter
Nooksack	Southfork	10GX	2010	adult	Sept - October	wild	summer
Nooksack	Southfork	11GO	2011	adult	August - October	wild	summer
Samish	Samish	08BN	2008	adult	Feb - April	wild	winter
Samish	Samish	12AP	2012	adult	Feb - March	wild	winter
Skagit	Cascade	12DA	2012	adult	May	wild	winter
Skagit	Finney Creek	10CQ	2010	adult	March - May	wild	winter
Skagit	Finney Creek	11BK	2011	adult	March - May	wild	winter
Skagit	Finney Creek	12FT	2012	adult	November	wild	summer
Skagit	Suiattle	10AQ	2010	adult	March - April	wild	winter
Skagit	Suiattle	11BM	2011	adult	April	wild	winter
Skagit	upper Skagit	08DQ	2008	adult	Feb - May	wild	winter
Skagit	upper Skagit	09BN	2009	adult	April	wild	winter
Skagit	upper Skagit	10AO	2010	adult	March - May	wild	winter
Skagit	upper Skagit	11BI	2011	adult	April - May	wild	winter
Skagit	upper Skagit	10NI	2010	adult	Nov 2010 - Jan 2011	wild	winter
Skagit	Nookachamps	12AO	2012	Juv. (adult = 2)	March, May	wild	winter
Skagit	Sauk	09DU	2009	adult	March - April	wild	winter
Skagit	Sauk	10AR	2010	adult	Feb - May	wild	winter
Skagit	Sauk	11BN	2011	adult	April - May	wild	winter
Stillaguamish	Canyon Creek	13KA	2013	juvenile	October	wild	summer
Stillaguamish	Deer Creek	95CG	1995	juvenile	unknown	wild	summer
Stillaguamish	Deer Creek	12FL	2012	adult	July	wild	summer
Stillaguamish	Deer Creek	13GE	2013	adult	October	wild	summer
Stillaguamish	Deer Creek	13KB	2013	juvenile	Sept - October	wild	summer
Stillaguamish	mixed	06BY	2006	smolt	unknown	wild	mixed
Snohomish	NF Skykomish	04HN	2004	juvenile	unknown	wild	summer
Snohomish	NF Skykomish	12FK	2012	adult	August - September	wild	summer
Snohomish	NF Skykomish	13GF	2013	adult	July - August	wild	summer
Snohomish	NF Skykomish	13LJ	2013	juvenile	October	wild	summer
Snohomish	Pilchuck River	12MN	2012	adult	Feb - April	wild	winter
Snohomish	Skykomish mainstem	13GH	2013	adult	Feb - April	wild	winter
Snohomish	NF Tolt (Snoqualmie)	11IW	2011	juvenile	September	wild	summer
Snohomish	NF Tolt (Snoqualmie)	12IS	2012	juvenile	September	wild	winter
Snohomish	SF Tolt (Snoqualmie)	10IX	2010	juvenile	September	wild	winter
Snohomish	Snoqualmie	13BC	2013	adult	Feb - April	wild	winter
Snohomish	SF Tolt (Snoqualmie)	10IW	2010	juvenile	September	wild	summer
Green	Mainstem	04AY	2004	adult	unknown	wild	winter
Green	Mainstem	07CO	2007	smolt	unknown	wild	winter
Green	Mainstem	08EF	2008	smolt	May - June	wild	winter
Green	Soos Creek	13EH	2013	adult	March - April	wild	winter

Table 1b. Collection data, sorted in same order as in Table 1a. Collections were aggregated into Operational Units (OUs), which were the primary units for analysis. OUs were aggregated into Demographically Independent Populations (DIPs; PSSTRT 2013), which are the primary management units. Total N is the number of samples genotyped per OU. Samples were removed if they were missing more than one-third of loci, showed at least one cutthroat allele, or if they showed a relatedness of half-sib/first cousin or closer to at least one other sample.

Code	PSSTRT DIP	Operational Unit	Total N	Removed Missing Loci	Removed Cut. Alleles	Removed Relatedness	Final N
11NW	Nooksack R Winter-Run	MainstemNookEarlyAd	24	0	0	1	23
12MP	Nooksack R Winter-Run	MainstemNookEarlyAd	22	0	1	0	21
13GC	Nooksack R Winter-Run	MainstemNookEarlyAd	12	0	0	0	12
12MQ	Nooksack R Winter-Run	NFNooksackAd	50	0	0	6	44
09MN	Nooksack R Winter-Run	NFNooksackJuv	61	0	23	20	18
10PY	Nooksack R Winter-Run	NFNooksackJuv	2	1	0	0	1
12CF	Nooksack R Winter-Run	SFNooksackWinterAd	42	0	0	1	41
10GX	South Fork Nooksack R Summer-Run	SFNooksackSummerAd	36	0	0	5	31
11GO	South Fork Nooksack R Summer-Run	SFNooksackSummerAd	31	1	0	2	28
08BN	Samish R Winter-Run	SamishRiver	42	1	0	9	32
12AP	Samish R Winter-Run	SamishRiver	46	0	0	5	41
12DA	Mainstem Skagit R Summer- and Winter-Run	CascadeRiverwinteradultSTHD	13	0	0	0	13
10CQ	Mainstem Skagit R Summer- and Winter-Run	FinneyCreekAdults	23	0	0	2	21
11BK	Mainstem Skagit R Summer- and Winter-Run	FinneyCreekAdults	30	0	0	2	28
12FT	Mainstem Skagit R Summer- and Winter-Run	FinneyCreeksummerSTHD	26	0	0	4	22
10AQ	Sauk R Summer- and Winter-Run	SuiattleAdults	17	0	0	1	16
11BM	Sauk R Summer- and Winter-Run	SuiattleAdults	34	0	0	3	31
08DQ	Mainstem Skagit R Summer- and Winter-Run	upperSkagitRiverAdults	20	1	0	0	19
09BN	Mainstem Skagit R Summer- and Winter-Run	upperSkagitRiverAdults	10	0	0	0	10
10AO	Mainstem Skagit R Summer- and Winter-Run	upperSkagitRiverAdults	25	0	0	2	23
11BI	Mainstem Skagit R Summer- and Winter-Run	upperSkagitRiverAdults	34	0	2	2	30
10NI	Mainstem Skagit R Summer- and Winter-Run	upperSkagitRiverlargeresidentOmykiss	8	0	1	0	7
12AO	Nookachamps Creek Winter-Run	NookachampsCreekjuvenileOmykiss	50	1	4	4	41
09DU	Sauk R Summer- and Winter-Run	SaukRiver	17	0	0	1	16
10AR	Sauk R Summer- and Winter-Run	SaukRiver	24	2	0	1	21
11BN	Sauk R Summer- and Winter-Run	SaukRiver	24	0	0	1	23
13KA	Canyon Creek Summer-Run	CanyonCreekSummerJuv	100	2	2	47	49
95CG	Deer Creek Summer-Run	DeerCreekJuveniles95	48	0	0	23	25
12FL	Deer Creek Summer-Run	DeerCreekSummerAdult	1	0	0	0	1
13GE	Deer Creek Summer-Run	DeerCreekSummerAdult	7	0	0	0	7
13KB	Deer Creek Summer-Run	DeerCreekSummerJuv13	101	0	0	38	63
06BY	NA (sample is aggregate)	StillaguamishRiverSmoltTrap	94	1	5	9	79
04HN	North Fork Skykomish Summer-Run	NFSkyJuv04	47	11	1	10	25
12FK	North Fork Skykomish Summer-Run	NFSkySumAd1213	10	0	0	0	10
13GF	North Fork Skykomish Summer-Run	NFSkySumAd1213	4	0	0	0	4
13LJ	North Fork Skykomish Summer-Run	NFSkySumJuv2013	100	2	2	32	64
12MN	Pilchuck R Winter-Run	PilchuckR12	50	1	0	3	46
13GH	Snohomish / Skykomish R Winter-Run	SkyWinAd13	21	0	0	0	21
11IW	Snoqualmie River Winter-Run	NFToltAboveJuv11	25	0	0	11	14
12IS	Snoqualmie River Winter-Run	NFToltBelowJuv11	50	0	1	7	42
10IX	Snoqualmie River Winter-Run	SFToltBelowJuv10	75	6	1	22	46
13BC	Snoqualmie River Winter-Run	SnoqualmieWinAd13	24	0	0	0	24
10IW	Tolt River Summer-Run	SFToltAboveJuv10	75	0	1	37	37
04AY	Green River Winter-Run	GreenR04	49	0	1	11	37
07CO	Green River Winter-Run	GreenR07	39	4	3	1	31
08EF	Green River Winter-Run	GreenR08	54	0	0	2	52
13EH	Green River Winter-Run	GreenRWildWinterBroodstock13	31	0	0	3	28

Table 2. Geographic and temporal scope, biological and management descriptors, and sample size of hatchery-origin collections used in this study. All samples were collected from segregated hatchery programs, either early winter (i.e., Chambers Creek – origin), or early summer (i.e., Skamania – origin), designated here as Operational Units, which were the primary units for analysis. Total N is the number of samples genotyped per OU. Samples were removed if they were missing more than one-third of loci, showed at least one cutthroat allele, or if they showed a relatedness of half-sib/first cousin or closer to at least one other sample. Final N was used for remaining analyses.

Basin	Hatchery/Program	Code	Collection Year	Life Stage	Origin	Program Type	OperationalUnit	Total N	Removed Poor Genotype	Removed Cut. Alleles	Removed Relatedness	Final N
Nooksack	Kendall Creek - early winter	01GA	2001	broodstock	hatchery	segregated	Kendall	100	0	0	60	40
Skagit	Marblemount - early winter	08LF	2008	broodstock	hatchery	segregated	MarblemountHatcheryAdults	46	0	4	17	25
Skagit	Marblemount - early winter	09CF	2009	broodstock	hatchery	segregated	MarblemountHatcheryAdults	56	0	0	13	43
Skagit	Marblemount - early winter	10AN	2010	broodstock	hatchery	segregated	MarblemountHatcheryAdults	50	0	0	21	29
Snohomish	Reiter Ponds - early summer	01GG	2001	broodstock	hatchery	segregated	ReiterPonds	39	0	0	4	35
Snohomish	Tokul Creek - early winter	01GC	2001	broodstock	hatchery	segregated	TokulHatchery	40	0	0	5	35
Green	Soos Creek - early winter	03LZ	2003	broodstock	hatchery	segregated	SoosChambers03	45	0	0	15	30
Green	Soos Creek - early summer	03MA	2003	broodstock	hatchery	segregated	SoosSkamania03	90	0	0	48	42

Table 3. SNP loci, with WDFW identifier, assay names, and reference for locus – source. Samples were genotyped using all loci. Loci were removed from analyses for a variety of reasons (see text). A check mark indicated that that locus was used for all analyses in that basin.

WDFW Identifier	Assay name	Reference	Database-wide Status	Genotyped In:							
				Green	Snohomish	Pilchuck	Stillaguamish	Skagit	Samish	Nooksack	Cross-Basin
AOmy005	Omy_aspAT-123	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy010	Omy_CRB2677.106	13	Omykiss Genotyping	✓	✓	✓	✓	✓	✓		
AOmy014	Omy_e1-147	13	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy015	Omy_gdh-271	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy016	Omy_GH1P1_2	2	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy021	Omy_LDHB-2_e5	2	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy023	Omy_MYC_2	2	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy026	Omy_myoD.178	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy027	Omy_nkef-241	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy028	Omy_nramp-146	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy029	Omy_Ogo4.212	4	Omykiss Genotyping			✓	✓	✓	✓		
AOmy042	Omy_BAC-F5.284	9	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy047	Omy_u07-79-166	9	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy048	Omy_113490-159	1	Omykiss Genotyping	✓				✓	✓	✓	
AOmy049	Omy_114315-438	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy051	Omy_121713-115	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy056	Omy_128693-455	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy058	Omy_130524-160	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy059	Omy_187760-385	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy061	Omy_96222-125	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓		
AOmy062	Omy_97077-73	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy065	Omy_97954-618	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy067	Omy_aromat-280	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy068	Omy_arp-630	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy072	Omy_cd59b-112	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy073	Omy_colla1-525	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy074	Omy_cox2-335	16	Removed - too few individuals scored								
AOmy078	Omy_g1-103	14	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy079	Omy_g12-82	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy081	Omy_gh-475	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy082	Omy_gsdf-291	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy084	Omy_hsc715-80	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy087	Omy_hsp47-86	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy088	Omy_hsp70aPro-329	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy089	Omy_hsp90BA-193	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy091	Omy_IL17-185	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy092	Omy_IL1b-163	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy094	Omy_inos-97	16	Removed - No Variation								
AOmy095	Omy_mapK3-103	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy096	Omy_mcsf-268	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy100	Omy_nach-200	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy105	Omy_OmyP9-180	13	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓

Table 3. Continued

WDFW Identifier	Assay name	Reference	Database-wide Status	Genotyped In:							
				Green	Snohomish	Pilchuck	Stillaguamish	Skagit	Samish	Nooksack	Cross-Basin
AOmy107	Omy_Ots249-227	4	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy108	Omy_oxct-85	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy110	Omy_star-206	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy111	Omy_stat3-273	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy113	Omy_tlr3-377	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy114	Omy_tlr5-205	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy117	Omy_u09-52-284	9	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy118	Omy_u09-53-469	9	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy120	Omy_u09-54.311	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy123	Omy_u09-55-233	9	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy125	Omy_u09-56-119	9	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy129	Omy_BAMBI4-238	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy132	Omy_G3PD_2.246	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy134	Omy_Il-1b-028	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy137	Omy_u09-61.043	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy144	Omy_UT16_2.173	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy147	Omy_U11_2b.154	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy149	Omy_gluR-79	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy152	Omy_SECC22b-88	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy173	BH2VHSVip10	11	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy174	OMS00003	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy176	OMS00013	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy177	OMS00018	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy179	OMS00041	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy180	OMS00048	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy181	OMS00052	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy182	OMS00053	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy183	OMS00056	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy184	OMS00057	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy185	OMS00061	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy186	OMS00062	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy187	OMS00064	12	Omykiss Genotyping		✓		✓			✓	
AOmy189	OMS00071	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy190	OMS00072	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy191	OMS00078	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy192	OMS00087	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy193	OMS00089	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy194	OMS00090	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy195	OMS00092	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy197	OMS00103	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy198	OMS00105	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy199	OMS00112	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy200	OMS00116	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy201	OMS00118	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓

Table 3. Continued

WDFW Identifier	Assay name	Reference	Database-wide Status	Genotyped In:							
				Green	Snohomish	Pilchuck	Stillaguamish	Skagit	Samish	Nooksack	Cross-Basin
AOmy202	OMS00119	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy203	OMS00120	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy204	OMS00121	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy205	OMS00127	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy206	OMS00128	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy207	OMS00132	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy208	OMS00133	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy209	OMS00134	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy210	OMS00153	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy211	OMS00154	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy212	OMS00156	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy213	OMS00164	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy214	OMS00169	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy215	OMS00175	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy216	OMS00176	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy218	OMS00180	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy220	Omy_1004	8	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy221	Omy_101554-306	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy222	Omy_101832-195	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy223	Omy_101993-189	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy225	Omy_102505-102	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy226	Omy_102867-443	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy227	Omy_103705-558	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy228	Omy_104519-624	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy229	Omy_104569-114	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy230	Omy_105075-162	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy231	Omy_105385-406	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy232	Omy_105714-265	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy233	Omy_107031-704	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy234	Omy_107285-69	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy235	Omy_107336-170	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy237	Omy_107806-34	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy238	Omy_108007-193	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy239	Omy_109243-222	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy240	Omy_109525-403	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy241	Omy_110064-419	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy242	Omy_110078-294	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy243	Omy_110362-585	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy244	Omy_110689-148	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy246	Omy_111084-526	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy247	Omy_111383-51	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy248	Omy_111666-301	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy249	Omy_112301-202	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy250	Omy_112820-82	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓

Table 3. Continued

WDFW Identifier	Assay name	Reference	Database-wide Status	Genotyped In:							
				Green	Snohomish	Pilchuck	Stillaguamish	Skagit	Samish	Nooksack	Cross-Basin
AOmy252	Omy_114976-223	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy253	Omy_116733-349	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy254	Omy_116938-264	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy255	Omy_117259-96	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy256	Omy_117286-374	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy257	Omy_117370-400	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy258	Omy_117540-259	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy260	Omy_117815-81	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy261	Omy_118175-396	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy262	Omy_118205-116	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy263	Omy_118654-91	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy265	Omy_120255-332	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy266	Omy_128996-481	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy267	Omy_129870-756	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy268	Omy_131460-646	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy269	Omy_98683-165	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy270	Omy_cyp17-153	16	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy271	Omy_ftzf1-217	16	Omykiss Genotyping	✓	✓	✓	✓				
AOmy272	Omy_GHSR-121	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy273	Omy_metA-161	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy274	Omy_UBA3b	8	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy275	M09AAC.055	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy276	M09AAE-082	15	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy277	OMGH1PROM1-SNP1	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy279	OMS00015	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy280	OMS00024	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy283	OMS00070	12	Omykiss Genotyping	✓				✓	✓	✓	
AOmy284	OMS00074	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy285	OMS00096	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy286	OMS00111	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy288	OMS00149	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy289	OMS00173	12	Removed - too few individuals scored								
AOmy290	Omy_105105-448	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy291	Omy_110201-359	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy292	Omy_128923-433	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy293	Omy_anp-17	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy294	Omy_bcAKala-380rd	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy295	Omy_cin-172	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy296	Omy_ndk-152	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy297	Omy_nips-299	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy298	Omy_ntl-27	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy299	Omy_rbm4b-203	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy300	Omy_sys1-188	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy301	Omy_txnip-343	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓

Table 3. Continued

WDFW Identifier	Assay name	Reference	Database-wide Status	Genotyped In:							
				Green	Snohomish	Pilchuck	Stillaguamish	Skagit	Samish	Nooksack	Cross-Basin
AOmy302	Omy_vamp5-303	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy303	Omy_vatf-406	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy305	OMS00077	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy306	OMS00101	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy311	Omy_G3PD_2-371	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy320	Omy_redd1-410	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy322	Omy_srp09-37	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy324	Omy1011	8	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy326	OMS00068	12	Omykiss Genotyping		✓	✓	✓	✓	✓	✓	
AOmy327	OMS00079	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy328	OMS00106	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy329	OMS00179	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy331	Omy_114587-480	1	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy335	OMS00017	12	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
AOmy341	Omy_metB-138	5	Omykiss Genotyping	✓	✓	✓	✓	✓	✓	✓	✓
ASpI001	Ocl_Okerca	10	Omykiss-Oclarki introgression ID only								
ASpI014	Omy_F5_136	6	Omykiss-Oclarki introgression ID only								
ASpI018	Omy_Omyclmk436-96	5	Omykiss-Oclarki introgression ID only								
			Total	183	183	183	184	184	184	182	178

1 Abadia-Cardoso et al. 2011
2 Aguilar and Garza 2008
3 Brunelli et al. 2008
4 Campbell and Narum 2009
5 CRITFC - N Campbell unpubl.
6 Finger et al. 2009
7 NOAA – JC Garza unpubl.
8 Hansen et al. 2011
9 Limborg et al. 2011
10 McGlauflin et al. 2010
11 UW – C Pascal and M Hansen unpubl.
12 Sánchez et al. 2009
13 Sprowles et al. 2006
14 Stephens et al. 2009
15 WDFW - S. Young unpubl.
16 WSU-J. DeKoning unpubl.

Table 4. The F_{ST} values from the Whidbey Basin (a) and Nooksack (b) modeled populations developed in the program *MS* (Hudson 2002). Data from each basin were divided into "Before" (all wild and hatchery samples from each basin), and "After" (only those samples included the phylogenetic and morphometric analyses; see "Identity of local assignments" in Methods). EWH = early winter hatchery. ESH = early summer hatchery.

(a)

Modeled Pop	Run #	EWH v Wild	ESH v Wild	EWH v ESH
Whidbey_E-2	1	0.032	0.064	0.061
Whidbey_E-2	3	0.027	0.049	0.051
Whidbey_E-2	4	0.031	0.065	0.059
Whidbey_E-2	5	0.024	0.050	0.046
Whidbey_E-2	6	0.027	0.061	0.061
Whidbey_E-2	10	0.022	0.044	0.046
Whidbey_E-2	11	0.031	0.056	0.056
Whidbey_E-2	15	0.033	0.054	0.054
Whidbey_E-2	16	0.030	0.050	0.049
Whidbey_E-2	17	0.027	0.053	0.056
Whidbey_E-2	**Mean**	**0.028**	**0.055**	**0.054**
Green	Before	0.025	0.051	0.024
Green	After	0.033	0.060	0.052
Snohomish	Before	0.032	0.045	0.053
Snohomish	After	0.035	0.072	0.055
Stillaguamish	Before	0.035	0.056	na
Stillaguamish	After	0.039	0.055	na
Skagit	Before	0.030	0.054	0.052
Skagit	After	0.036	0.061	0.052

(b)

Modeled Pop	Run #	EWH v Wild	ESH v Wild	EWH v ESH
Nooksack_v2_09	4	0.051	0.075	0.059
Nooksack_v2_09	21	0.056	0.080	0.060
Nooksack_v2_09	22	0.055	0.074	0.061
Nooksack_v2_09	23	0.059	0.078	0.060
Nooksack_v2_09	28	0.054	0.078	0.056
Nooksack_v2_09	30	0.044	0.070	0.049
Nooksack_v2_09	52	0.053	0.075	0.046
Nooksack_v2_09	63	0.048	0.078	0.057
Nooksack_v2_09	81	0.046	0.074	0.055
Nooksack_v2_09	**Mean**	**0.052**	**0.076**	**0.056**
Nooksack	Before	0.052	0.073	0.051
Nooksack	After	0.057	0.078	0.053

Table 5. Distribution of assignments from *Structure* and associated assignment error rates for the Whidbey Basin (upper) and Nooksack (lower) modeled populations. The *Structure* assigned categories are the rows and the source categories are the columns. Each source category consists of 1000 individuals; 100 individuals each from the ten model populations.

		Source Category								
	Assigned Category	Early Winter Hatchery (EWH)	Hybrid: EWH - Wild	Hybrid: ESH - Wild	Early Summer Hatchery (ESH)	Wild	Hybrid: EWH - ESH	Total Assigned	Total that should be assigned - No Call	Assignment Error Rate
Whidbey Basin Model	Early Winter Hatchery (EWH)	**655**	113	4	1	5	66	844	937	0.22
	Hybrid: EWH - Wild	216	**625**	58	1	229	33	1162	920	0.46
	Hybrid: ESH - Wild	0	11	**514**	24	26	37	612	782	0.16
	Early Summer Hatchery (ESH)	0	0	87	**839**	0	159	1085	963	0.23
	Wild	5	153	79	0	**693**	1	931	956	0.26
	Hybrid: EWH - ESH	61	18	40	98	3	**507**	727	803	0.30
	No Call	63	80	218	37	44	197	639		na
	Total Source	1000	1000	1000	1000	1000	1000	6000		0.29
	No Call Rate	0.06	0.08	0.22	0.04	0.04	0.20	0.11		
	Source Error Rate Total	0.35	0.38	0.49	0.16	0.31	0.49	0.36		
	Source Error Rate Assigned Only	0.30	0.32	0.34	0.13	0.28	0.37	0.29		
Nooksack Model	Early Winter Hatchery (EWH)	**771**	90	0	0	0	47	908	981	0.15
	Hybrid: EWH - Wild	134	**692**	15	0	126	13	980	913	0.29
	Hybrid: ESH - Wild	0	15	**673**	47	52	35	822	825	0.18
	Early Summer Hatchery (ESH)	0	0	69	**898**	0	76	1043	989	0.14
	Wild	1	94	32	0	**801**	0	928	979	0.14
	Hybrid: EWH - ESH	75	22	36	44	0	**656**	833	827	0.21
	No Call	19	87	175	11	21	173	486		na
	Total Source	1000	1000	1000	1000	1000	1000	6000		0.19
	No Call Rate	0.02	0.09	0.18	0.01	0.02	0.17	0.08		
	Source Error Rate Total	0.23	0.31	0.33	0.10	0.20	0.34	0.25		
	Source Error Rate Assigned Only	0.21	0.24	0.18	0.09	0.18	0.21	0.19		

Table 6a. The proportion of individuals from each Operational Unit and DIP (PSSHTRT 2013) assigned by *Structure* (Pritchard et al. 2000, Falush et al. 2003) analysis to specific lineage categories. Pure wild lineages are designated as Basin Winter (i.e., occurs within more than one subbasin), or Local Winter or Summer (i.e., occurs within only one creek or river within the basin); pure hatchery lineages are designated as either Early Winter Hatchery (Chambers Creek – origin) or Early Summer Hatchery (Skamania – origin); and mixed lineages between pure wild and pure hatchery lineages are designed as Hybrid. Below the category names, in italic typeface, are the model categories (see text and Figure 1) to which the categories here are aggregated. Hatchery-Lineage Weights are used for calculating effective pHOS (see text) and refer to the average number of hatchery lineages within each category, with pure wild = 0, hybrid = 1, and pure hatchery = 2. Only those categories that are a mix (hybrid) between pure hatchery and wild are designated as Hatchery-Wild Introgression.

(a) Green River Basin.

Population	N	Basin Winter	Early Winter Hatchery	Early Summer Hatchery	Hybrid Early Summer Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Early Summer Hatchery
		Wild	*EWH*	*ESH*	*Hybrid:ESH-Wild*	*Hybrid:EWH-Wild*	*Hybrid:EWH-ESH*
# Hatchery-Lineage Weight		0	2	2	1	1	1
Hatchery-Wild Introgression		No	No	No	Yes	Yes	No
Operational Unit							
GreenR04	36	1.00	0.00	0.00	0.00	0.00	0.00
GreenR07	31	0.73	0.01	0.00	0.00	0.24	0.02
GreenR08	52	0.73	0.00	0.02	0.00	0.25	0.01
GreenRWildWinterBroodstock13	27	1.00	0.00	0.00	0.00	0.00	0.00
SoosChambers03	30	0.00	0.68	0.32	0.00	0.00	0.00
SoosSkamania03	41	0.00	0.01	0.99	0.00	0.00	0.00
DIP - Adult samples only							
Green River Winter-Run	63	1.00	0.00	0.00	0.00	0.00	0.00
DIP - All samples							
Green River Winter-Run	146	0.88	0.00	0.01	0.00	0.11	0.00

Table 6b. Snohomish River Basin. See Table 6a for description.

Population	N	Basin Winter	Pilchuck Local Winter	Hybrid Basin Winter - Pilchuck Local Winter	Early Winter Hatchery	Early Summer Hatchery	Local Summer Hatchery Ancestry[1]	Hybrid Early Summer Hatchery - Local Summer Hatchery Ancestry	Hybrid Early Summer Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Early Summer Hatchery	Hybrid Early Winter Hatchery - Local Summer Hatchery Ancestry
		Wild	*Wild*	*Wild*	*EWH*	*ESH*	*ESH*	*ESH*	*Hybrid:ESH-Wild*	*Hybrid:EWH-Wild*	*Hybrid:EWH-ESH*	*Hybrid:EWH-ESH*
Hatchery-Lineage Weight		0	0	0	2	2	2	2	1	1	1	1
Hatchery-Wild Introgression		No	No	No	No	No	Yes	No	Yes	Yes	No	No
Operational Unit												
NFSkyJuv04	21	0.00	0.00	0.00	0.07	0.16	0.26	0.26	0.00	0.00	0.08	0.16
NFSkySumAd2013	11	0.00	0.00	0.00	0.00	0.18	0.64	0.18	0.00	0.00	0.00	0.00
NFSkySumJuv2013	57	0.00	0.00	0.00	0.00	0.09	0.58	0.33	0.00	0.00	0.00	0.00
NFToltAboveJuv11	7	0.74	0.00	0.00	0.00	0.16	0.00	0.00	0.10	0.00	0.00	0.00
NFToltBelowJuv11	39	0.60	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.37	0.00	0.00
PilchuckR12	40	0.50	0.28	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SFToltAboveJuv10	33	0.00	0.00	0.00	0.00	0.21	0.48	0.31	0.00	0.00	0.00	0.00
SFToltBelowJuv10	37	0.42	0.00	0.00	0.00	0.16	0.00	0.00	0.15	0.21	0.06	0.00
SkyWinAd13	20	0.76	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00
SnoqualmieWinAd13	24	0.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00
DIP - Adult samples only												
North Fork Skykomish Summer-Run	11	0.00	0.00	0.00	0.00	0.18	0.64	0.18	0.00	0.00	0.00	0.00
Tolt River Summer-Run	0	na	na	na	na	na	na	na	na	na	na	na
Snoqualmie River Winter-Run	24	0.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00
Snohomish / Skykomish R Winter-Run	20	0.76	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00
Pilchuck R Winter-Run	40	0.50	0.28	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DIP - All samples												
North Fork Skykomish Summer-Run	89	0.00	0.00	0.00	0.02	0.12	0.55	0.31	0.00	0.00	0.00	0.00
Tolt River Summer-Run	33	0.00	0.00	0.00	0.00	0.21	0.48	0.31	0.00	0.00	0.00	0.00
Snoqualmie River Winter-Run	113	0.71	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.28	0.00	0.00
Snohomish / Skykomish R Winter-Run	20	0.76	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00
Pilchuck R Winter-Run	40	0.50	0.28	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

[1] The phylogenetic relationship of this category is ambiguous (Figures 7, 8). As such we conservatively estimate that both parental lineages are ESH, but since there is a possibility that the category may have hybridized historically with a now potentially extirpated wild population, or other groups, we also scored this category as introgressed.

Table 6c. Stillaguamish River Basin. See Table 6a for description.

Population	N	Local Summer Canyon C	Local Summer Deer C	Hybrid Canyon C - Deer C	Early Summer Hatchery	Hybrid Early Summer Hatchery - Canyon C	Hybrid Early Summer Hatchery - Deer C	Hybrid Early Winter Hatchery - Early Summer Hatchery
		Wild	*Wild*	*Wild*	*ESH*	*Hybrid:ESH-Wild*	*Hybrid:ESH-Wild*	*Hybrid:EWH-ESH*
# Hatchery-Lineage Weight		0	0	0	2	1	1	1
Hatchery-Wild Introgression		No	No	No	No	Yes	Yes	No
Operational Unit								
CanyonCreekSummerJuv	44	0.69	0.00	0.31	0.00	0.00	0.00	0.00
DeerCreekJuveniles95	23	0.00	0.77	0.00	0.00	0.00	0.21	0.03
DeerCreekSummerAdult	8	0.00	0.77	0.00	0.00	0.00	0.23	0.00
DeerCreekSummerJuv13	59	0.02	0.82	0.16	0.00	0.00	0.00	0.00
StillaguamishRiverSmoltTrap	62	0.52	0.17	0.19	0.09	0.01	0.02	0.01
DIP - Adult samples only								
Canyon Creek Summer-Run	0	na	na	na	na	na	na	na
Deer Creek Summer-Run	8	0.00	0.77	0.00	0.00	0.00	0.23	0.00
DIP - All samples								
Canyon Creek Summer-Run	44	0.69	0.00	0.31	0.00	0.00	0.00	0.00
Deer Creek Summer-Run	90	0.01	0.85	0.11	0.00	0.00	0.02	0.00

Table 6d. Skagit River Basin. See Table 6a for description.

Category	N	Basin Winter	Nookachamps Local Winter	Hybrid Basin Winter - Nookachamps Local Winter	Finney Creek Local Summer	Hybrid Basin Winter - Finney Creek Local Summer	Early Winter Hatchery	Early Summer Hatchery	Hybrid Early Summer Hatchery - Finney Creek Local Summer	Hybrid Early Winter Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Early Summer Hatchery
		Wild	*Wild*	*Wild*	*Wild*	*Wild*	*EWH*	*ESH*	*Hybrid:ESH-Wild*	*Hybrid:EWH-Wild*	*Hybrid:EWH-ESH*
# Hatchery-Lineage Weight		0	0	0	0	0	2	2	1	1	1
Hatchery-Wild Introgression		No	No	No	No	No	No	No	Yes	Yes	No
Operational Unit											
CascadeRiverwinteradultSTHD	13	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00
FinneyCreekAdults	49	0.53	0.00	0.00	0.13	0.32	0.02	0.00	0.00	0.00	0.00
FinneyCreeksummerSTHD	18	0.00	0.00	0.00	0.32	0.57	0.00	0.00	0.12	0.00	0.00
NookachampsCreekjuvenileOmykiss	39	0.03	0.49	0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SaukRiver	60	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SuiattleAdults	47	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
upperSkagitRiverAdults	81	0.99	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
upperSkagitRiverlargeresidentOmykiss	7	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DIP - Adult samples only											
Mainstem Skagit R Summer- and Winter-Run	172	0.91	0.00	0.00	0.02	0.05	0.01	0.00	0.00	0.00	0.00
Sauk R Summer- and Winter-Run	107	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nookachamps Creek Winter-Run	0	na	na	na	na	na	na	na	na	na	na
DIP - All samples											
Mainstem Skagit R Summer- and Winter-Run	172	0.91	0.00	0.00	0.02	0.05	0.01	0.00	0.00	0.00	0.00
Sauk R Summer- and Winter-Run	107	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nookachamps Creek Winter-Run	39	0.03	0.49	0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6e. Samish River Basin. See Table 6a for description.

Category	N	Basin Winter	Early Winter Hatchery	Early Summer Hatchery	Hybrid Early Winter Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Early Summer Hatchery
		Wild	*EWH*	*ESH*	*Hybrid:EWH-Wild*	*Hybrid:EWH-ESH*
# Hatchery-Lineage Weight		0	2	2	1	1
Hatchery-Wild Introgression		No	No	No	Yes	No
Operational Unit						
SamishRiver	72	0.83	0.00	0.00	0.17	0.00
DIP - Adult samples only						
SamishRiver	72	0.83	0.00	0.00	0.17	0.00
DIP - All samples						
SamishRiver	72	0.83	0.00	0.00	0.17	0.00

Table 6f. Nooksack River Basin. See Table 6a for description.

Category	N	Basin Winter	Hybrid Basin Winter - Local Summer	Local Summer	Early Winter Hatchery	Hybrid Early Summer Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Basin Winter	Hybrid Early Winter Hatchery - Early Summer Hatchery
		Wild	*Wild*	*Wild*	*EWH*	*Hybrid:ESH-Wild*	*Hybrid:EWH-Wild*	*Hybrid:ESH-ESH*
# Hatchery-Lineage Weight		0	0	0	2	1	1	1
Hatchery-Wild Introgression		No	No	No	No	Yes	Yes	No
Operational Unit								
MainstemNookEarlyAd	55	0.86	0.00	0.00	0.02	0.03	0.08	0.01
NFNooksackAd	43	0.98	0.00	0.00	0.02	0.00	0.00	0.00
NFNooksackJuv	19	0.90	0.00	0.00	0.00	0.10	0.00	0.00
SFNooksackSummerAd	59	0.02	0.09	0.90	0.00	0.00	0.00	0.00
SFNooksackWinterAd	39	0.70	0.24	0.03	0.03	0.00	0.00	0.00
DIP - Adult samples only								
Nooksack R Winter-Run	137	0.91	0.06	0.01	0.02	0.00	0.00	0.00
South Fork Nooksack R Summer-Run	59	0.02	0.09	0.90	0.00	0.00	0.00	0.00
DIP - All samples								
Nooksack R Winter-Run	157	0.89	0.04	0.00	0.02	0.02	0.00	0.03
South Fork Nooksack R Summer-Run	59	0.02	0.09	0.90	0.00	0.00	0.00	0.00

Table 7. Operational Units' estimated spawning proportion within specific DIPs (A. Hoffmann, WDFW, pers. comm. 2014)

PSSTRT DIP	Operational Unit	Proportion of Spawning within DIP
Snoqualmie River Winter-Run	NFToltAbove&BelowJuv11	0.0411
Snoqualmie River Winter-Run	SFToltBelowJuv10	0.0589
Snoqualmie River Winter-Run	SnoqualmieWinAd13	0.9000
Mainstem Skagit R Summer- and Winter-Run	CascadeRiverwinteradultSTHD	0.1194
Mainstem Skagit R Summer- and Winter-Run	upperSkagitRiverAdults	0.7384
Mainstem Skagit R Summer- and Winter-Run	FinneyCreekSummer&WinterAdults	0.1422
Sauk R Summer- and Winter-Run	SaukRiver	0.7149
Sauk R Summer- and Winter-Run	SuiattleAdults	0.2851
Nooksack R Winter-Run	MainstemNookEarlyAd	0.2754
Nooksack R Winter-Run	SFNooksackWinterAd	0.2484
Nooksack R Winter-Run	NFNooksackAd&Juv	0.4762

Table 8a. Early Winter Hatchery (Chambers Creek), Early Summer Hatchery (Skamania), and Total effective pHOS, and Introgression for each Operational Unit and DIP. Introgression is the sum of the proportions assigned to categories identified as Hatchery-Wild Introgression, for each Operational Unit and TRT DIP (see Table 6). See text for calculation of effective pHOS.

(a) Green River Basin.

Population	Effective pHOS Early Winter	Effective pHOS Early Summer	Effective pHOS Total	Introgression
Operational Unit				
GreenR04	0.00	0.00	0.00	0.00
GreenR07	0.14	0.01	0.15	0.24
GreenR08	0.13	0.02	0.15	0.25
GreenRWildWinterBroodstock13	0.00	0.00	0.00	0.00
DIP - Adult samples only				
Green River Winter-Run	0.00	0.00	0.00	0.00
DIP - All samples				
Green River Winter-Run	0.05	0.01	0.06	0.11

Table 8b. Snohomish River Basin. See Table 8a for description.

Population	Effective pHOS Early Winter	Effective pHOS Early Summer	Effective pHOS Total	Introgression
Operational Unit				
NFSkyJuv04	0.19	0.81	1.00	0.26
NFSkySumAd2013	0.00	1.00	1.00	0.64
NFSkySumJuv2013	0.00	1.00	1.00	0.58
NFToltAboveJuv11	0.00	0.21	0.21	0.10
NFToltBelowJuv11	0.19	0.03	0.21	0.37
PilchuckR12	0.00	0.00	0.00	0.00
SFToltAboveJuv10	0.00	1.00	1.00	0.48
SFToltBelowJuv10	0.13	0.26	0.40	0.36
SkyWinAd13	0.00	0.24	0.24	0.00
SnoqualmieWinAd13	0.12	0.00	0.12	0.24
DIP - Adult samples only				
North Fork Skykomish Summer-Run	0.00	1.00	1.00	0.64
Tolt River Summer-Run	na	na	na	na
Snoqualmie River Winter-Run	0.12	0.00	0.12	0.24
Snohomish / Skykomish R Winter-Run	0.00	0.24	0.24	0.00
Pilchuck R Winter-Run	0.00	0.00	0.00	0.00
DIP - All samples				
North Fork Skykomish Summer-Run	0.02	0.98	1.00	0.55
Tolt River Summer-Run	0.00	1.00	1.00	0.48
Snoqualmie River Winter-Run	0.13	0.02	0.15	0.27
Snohomish / Skykomish R Winter-Run	0.00	0.24	0.24	0.00
Pilchuck R Winter-Run	0.00	0.00	0.00	0.00

Table 8c. Stillaguamish River Basin. See Table 8a for description.

Population	Effective pHOS Early Winter	Effective pHOS Early Summer	Effective pHOS Total	Introgression
Operational Unit				
CanyonCreekSummerJuv	0.00	0.00	0.00	0.00
DeerCreekJuveniles95	0.01	0.12	0.13	0.21
DeerCreekSummerAdult	0.00	0.11	0.11	0.23
DeerCreekSummerJuv13	0.00	0.00	0.00	0.00
StillaguamishRiverSmoltTrap	0.00	0.11	0.12	0.03
DIP - Adult samples only				
Canyon Creek Summer-Run	na	na	na	na
Deer Creek Summer-Run	0.00	0.11	0.11	0.23
DIP - All samples				
Canyon Creek Summer-Run	0.00	0.00	0.00	0.00
Deer Creek Summer-Run	0.00	0.01	0.01	0.02

Table 8d. Skagit River Basin. See Table 8a for description.

Population	Effective pHOS Early Winter	Effective pHOS Early Summer	Effective pHOS Total	Introgression
Operational Unit				
CascadeRiverwinteradultSTHD	0.08	0.00	0.08	0.17
FinneyCreekAdults	0.02	0.00	0.02	0.00
FinneyCreeksummerSTHD	0.00	0.06	0.06	0.12
NookachampsCreekjuvenileOmykiss	0.00	0.00	0.00	0.00
SaukRiver	0.00	0.00	0.00	0.00
SuiattleAdults	0.00	0.00	0.00	0.00
upperSkagitRiverAdults	0.01	0.00	0.01	0.00
upperSkagitRiverlargeresidentOmykiss	0.00	0.00	0.00	0.00
DIP - Adult samples only				
Mainstem Skagit R Summer- and Winter-Run	0.01	0.00	0.01	0.00
Sauk R Summer- and Winter-Run	0.00	0.00	0.00	0.00
Nookachamps Creek Winter-Run	na	na	na	na
DIP - All samples				
Mainstem Skagit R Summer- and Winter-Run	0.01	0.00	0.01	0.00
Sauk R Summer- and Winter-Run	0.00	0.00	0.00	0.00
Nookachamps Creek Winter-Run	0.00	0.00	0.00	0.00

Table 8e. Samish River Basin. See Table 8a for description.

Population	Effective pHOS Early Winter	Effective pHOS Early Summer	Effective pHOS Total	Introgression
Operational Unit				
SamishRiver	0.08	0.00	0.08	0.17
DIP - Adult samples only				
SamishRiver	0.08	0.00	0.08	0.17
DIP - All samples				
SamishRiver	0.08	0.00	0.08	0.17

Table 8f. Nooksack River Basin. See Table 8a for description.

Population	Effective pHOS Early Winter	Effective pHOS Early Summer	Effective pHOS Total	Introgression
Operational Unit				
MainstemNookEarlyAd	0.07	0.02	0.09	0.11
NFNooksackAd	0.02	0.00	0.02	0.00
NFNooksackJuv	0.00	0.05	0.05	0.10
SFNooksackSummerAd	0.00	0.00	0.00	0.00
SFNooksackWinterAd	0.03	0.00	0.03	0.00
DIP - Adult samples only				
Nooksack R Winter-Run	0.02	0.00	0.02	0.00
South Fork Nooksack R Summer-Run	0.00	0.00	0.00	0.00
DIP - All samples				
Nooksack R Winter-Run	0.03	0.02	0.05	0.02
South Fork Nooksack R Summer-Run	0.00	0.00	0.00	0.00

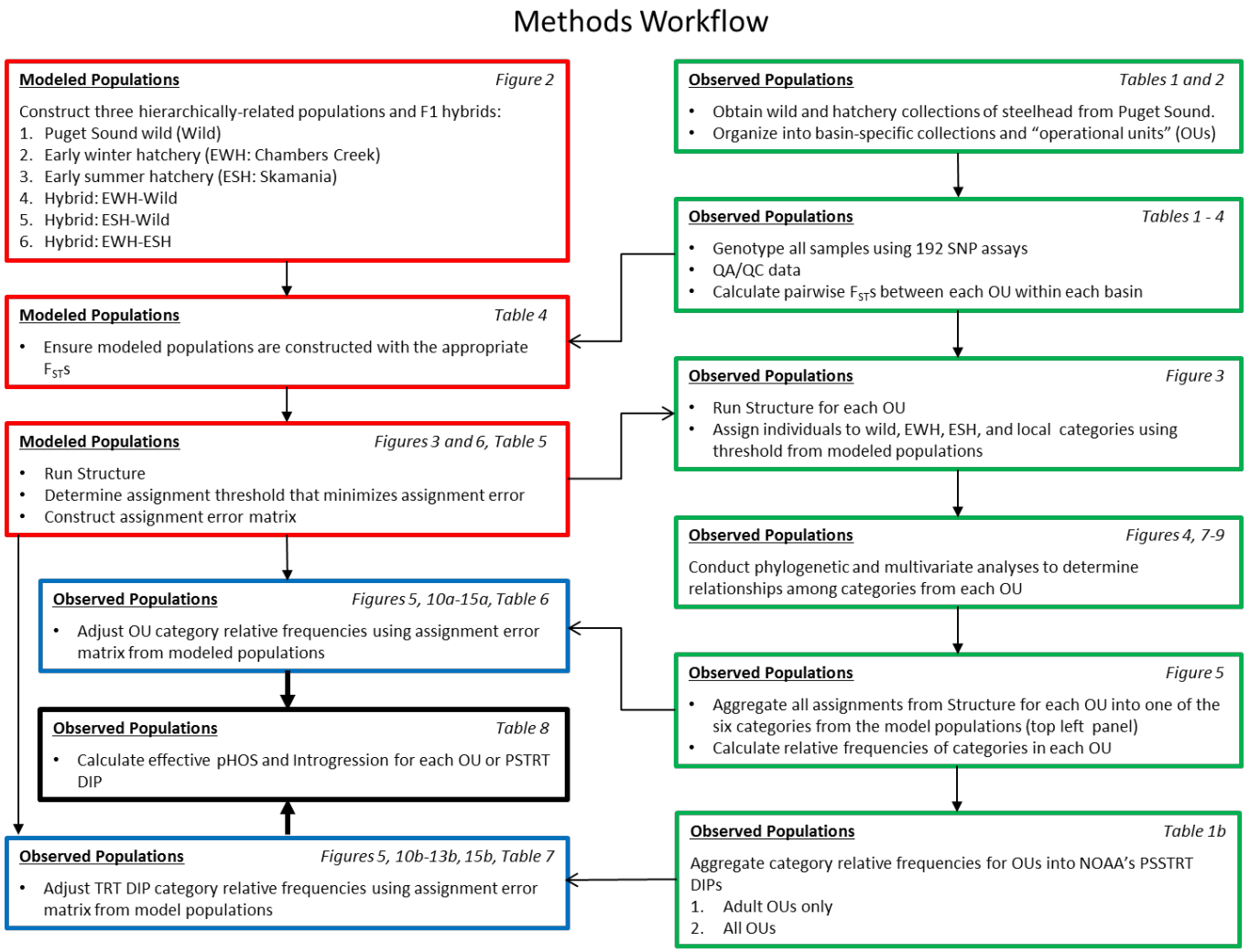

Figure 1. Basic design of workflow and index of methods and results for an analysis pipeline to produce estimates of effective pHOS and introgression within each Operational Unit (OU) and demographically independent population (DIP). The pipeline itself has four parts: (1) assignments of empirical data to categories (six green boxes to the right), (2) construction and assignments of modeled populations to categories, and determination of assignment error rates (three red boxes to the upper left); (3) adjustments to empirical data assignments to account for assignment errors (two blue boxes to lower left); and (4) characterization of hatcheries' genetic effects on wild populations using effective pHOS and introgression (black box to lower left).

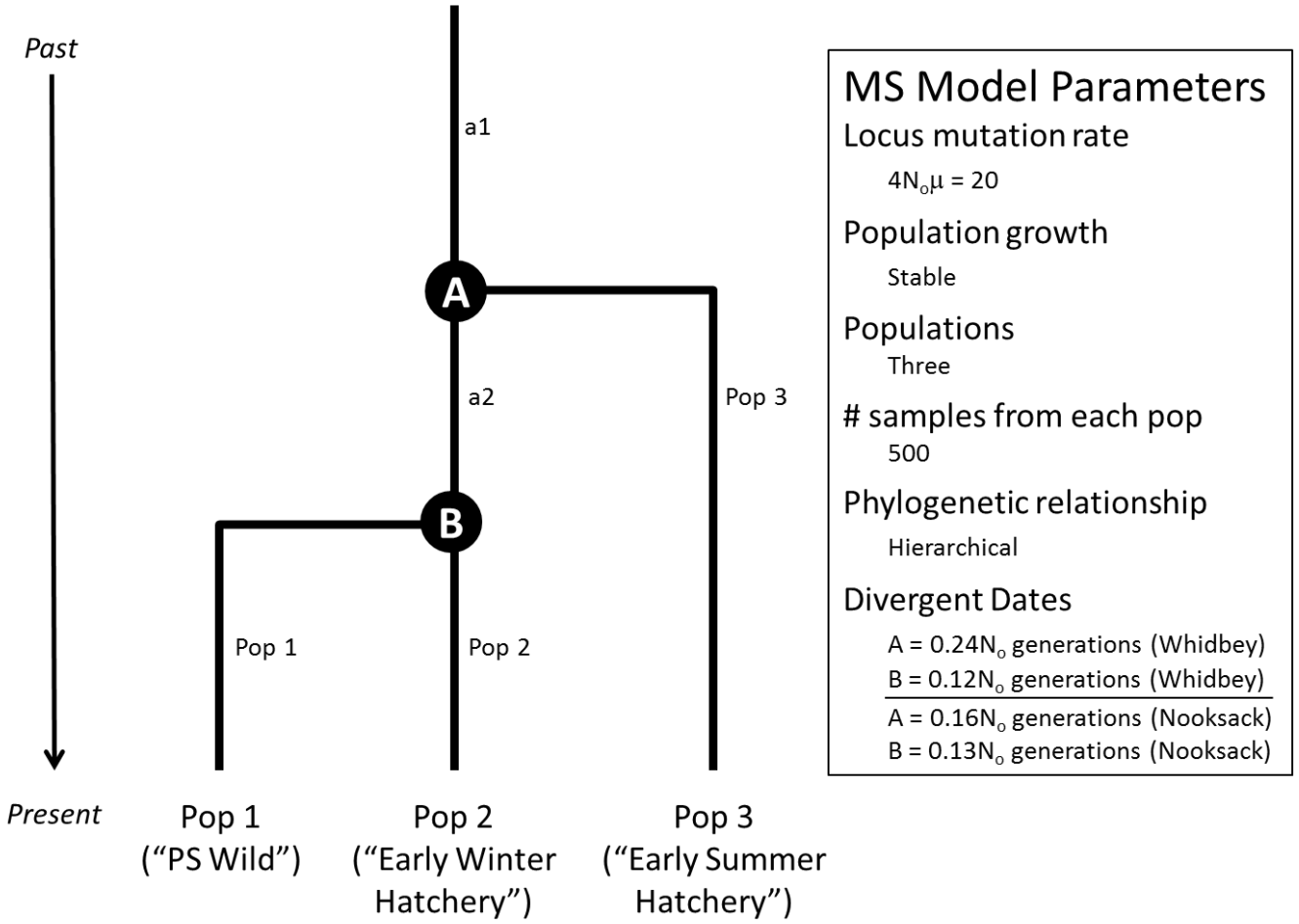

Figure 2. Schema and parameters used to develop modeled populations in the program *MS* (Hudson 2002). Modeled populations were hierarchically related. Single population existed at some time in the past (Pop a1). At time A, Pop a1 instantaneously split into two populations (Pop a2 and Pop 3) of the same size. The two populations were demographically stable and exchanged no immigrants. At time B, Pop a2 instantaneously split into two populations (Pop 1 and Pop 2) of the same size. The two populations were demographically stable and exchanged no immigrants. N_o = diploid population size and μ = neutral mutation rate. This represents the assumed relationship among Puget Sound wild, early winter hatchery and early summer hatchery populations. The locus mutation rate and divergent dates were determined by trial and error and were designed so that the divergence among Pop 1, 2, and 3 matched the empirical divergence among wild, early winter hatchery, and early summer hatchery steelhead collections in Puget Sound, respectively (see Table 4). Different divergence dates were needed for the Whidbey Basin and Nooksack models.

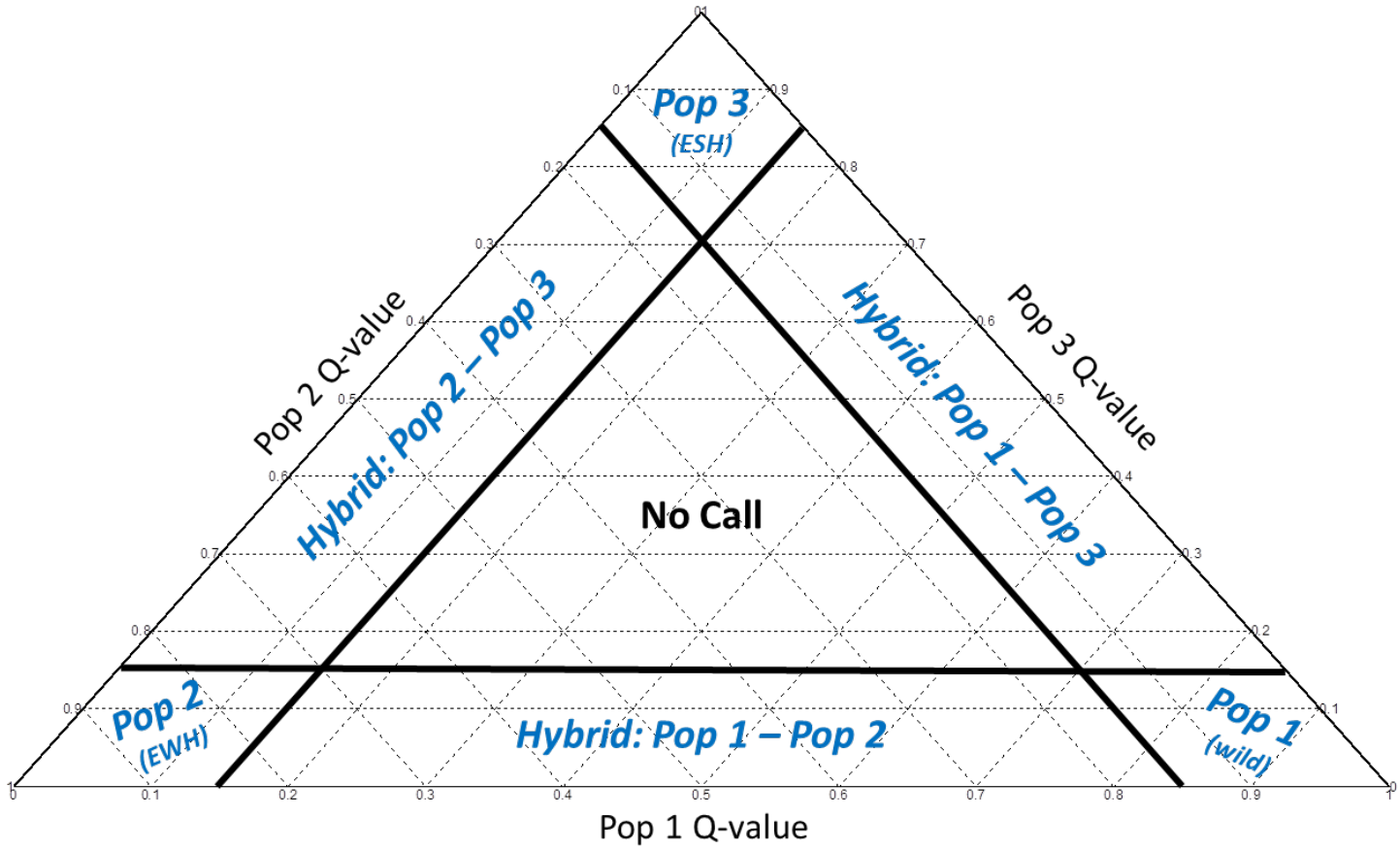

Figure 3. Ternary diagram indicating Structure k = 3 assignment regions and thresholds. Assignment thresholds (thick solid black lines) were set at Q-score = 0.15. See Figure 2 and text.

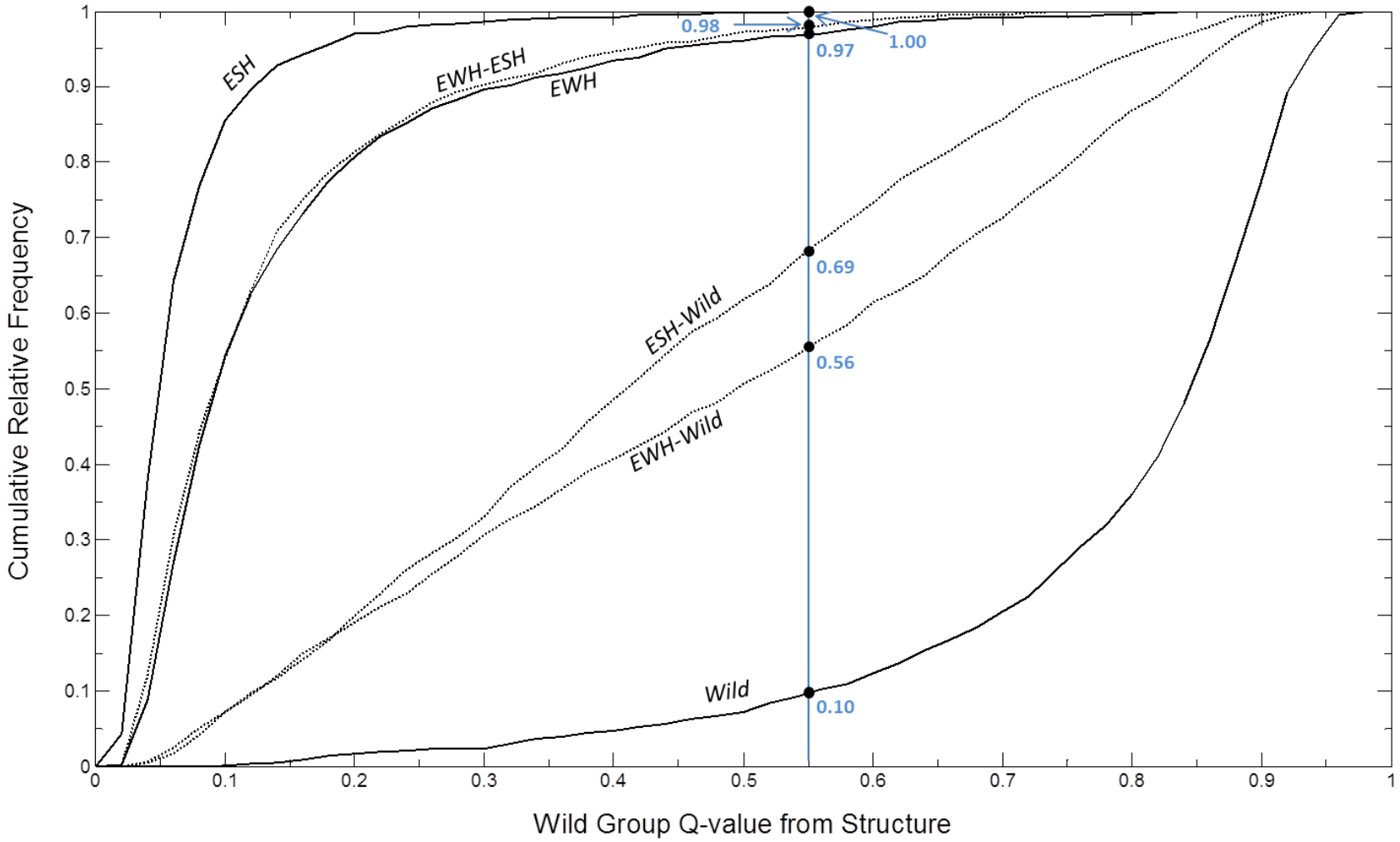

Figure 4. Cumulative frequency distribution of *Structure* Q-scores for the wild assigned category in the Whidbey Basin modeled populations, for the Wild, Hybrid:EWH-Wild, Hybrid:ESH-Wild, EWH, Hybrid:EWH-ESH, and ESH source categories. For multivariate and phylogenetic analyses, we used a Wild or Local assigned category Q-score ≥ 0.55 (area to right of vertical blue line). Using this threshold, we expected to include 90% of the Wild, but also 44%, 31%, 3% 2%, 0% of the Hybrid:EWH-Wild, Hybrid:ESH-Wild, EWH, Hybrid:EWH-ESH, and ESH source category individuals. EWH = early winter hatchery, ESH = early summer hatchery.

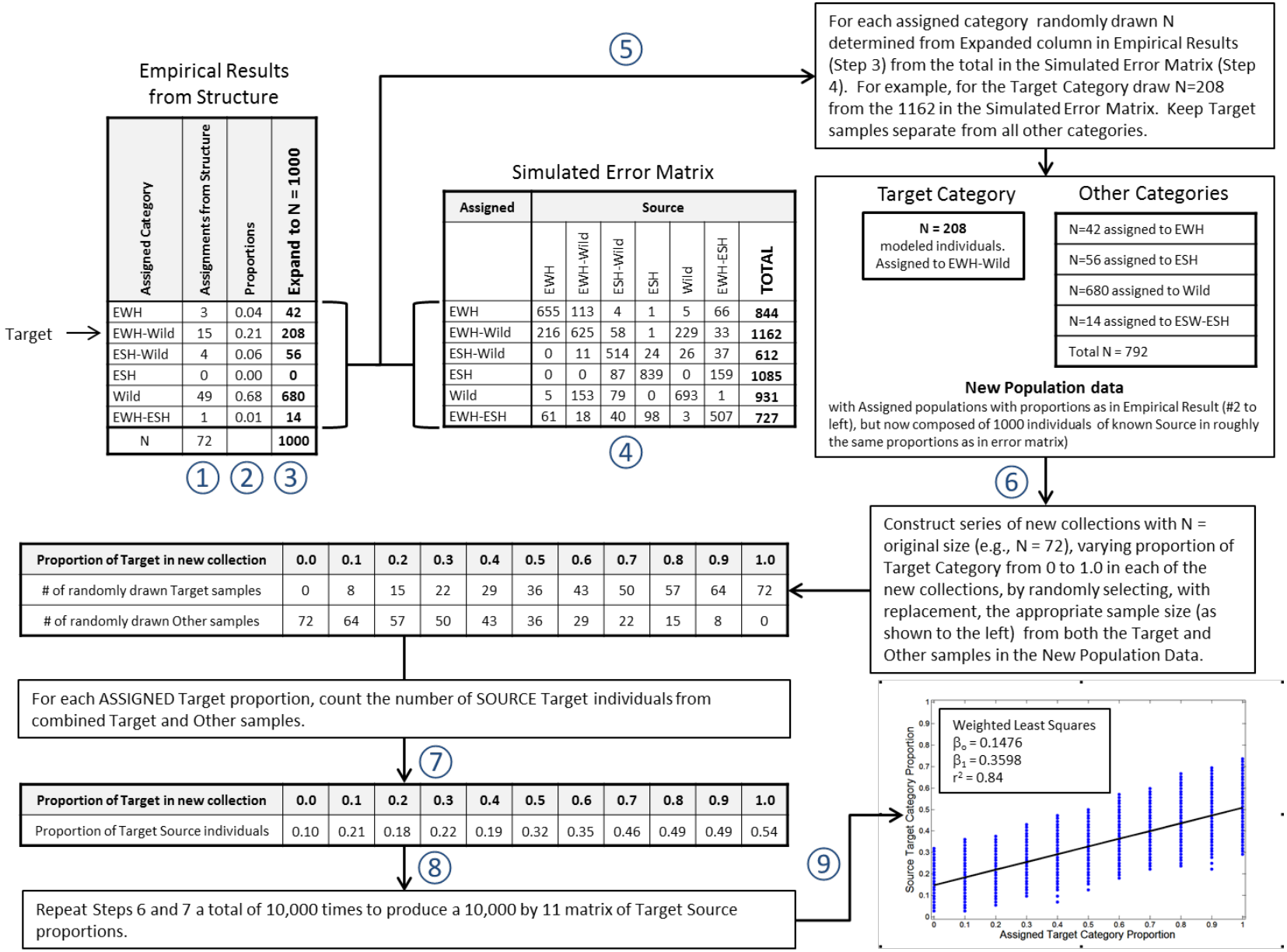

Figure 5. Likelihood-based procedure to correct *Structure* results to account for close phylogenetic relationships between the hatchery populations and wild populations. Numbered circles are procedure steps explained more fully in the text.

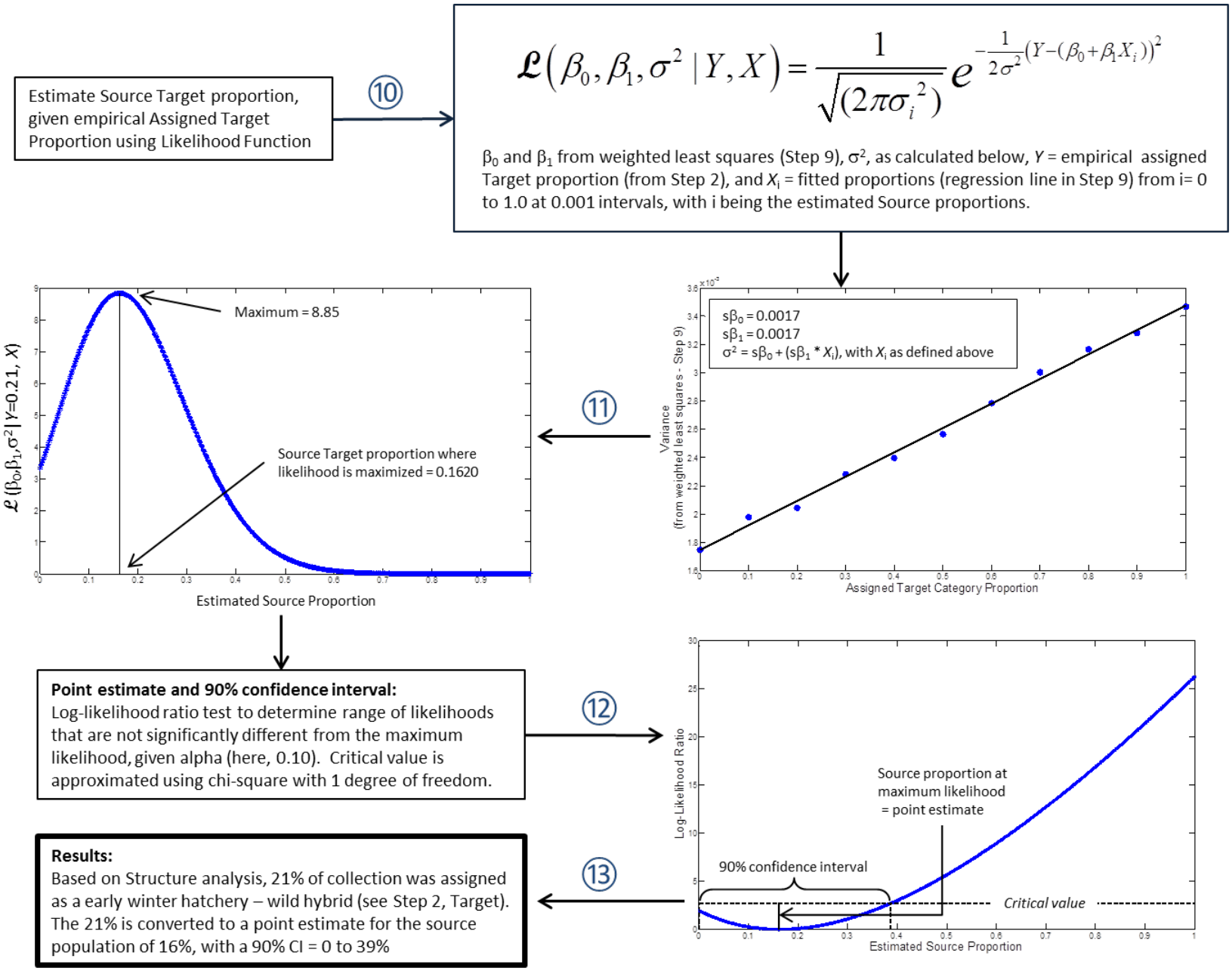

Figure 5. Continued.

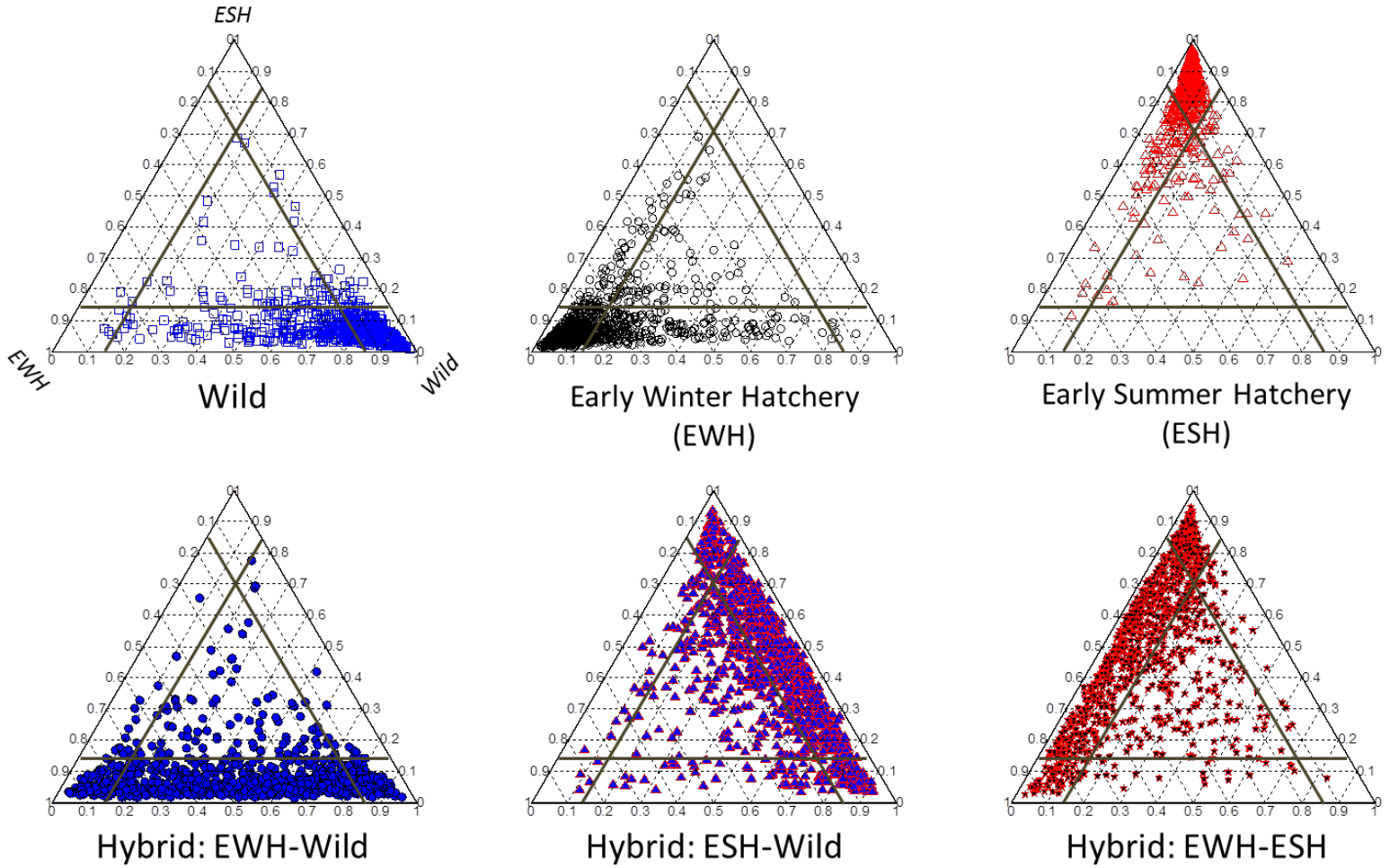

Figure 6. Distribution of the Wild (bottom axis), EWH (left axis), and ESH (right axis) assigned category Q-scores from k = 3 *Structure* analysis of the Whidbey Basin modeled populations, for the six source categories (one ternary plot per source category). The actual assignments using threshold value = 0.15 (solid black lines) for the source categories are indicated in the upper portion of Table 5. The highest density of Q-scores for each source category occurred in the appropriate assignment region for that source category (see Figure 3).

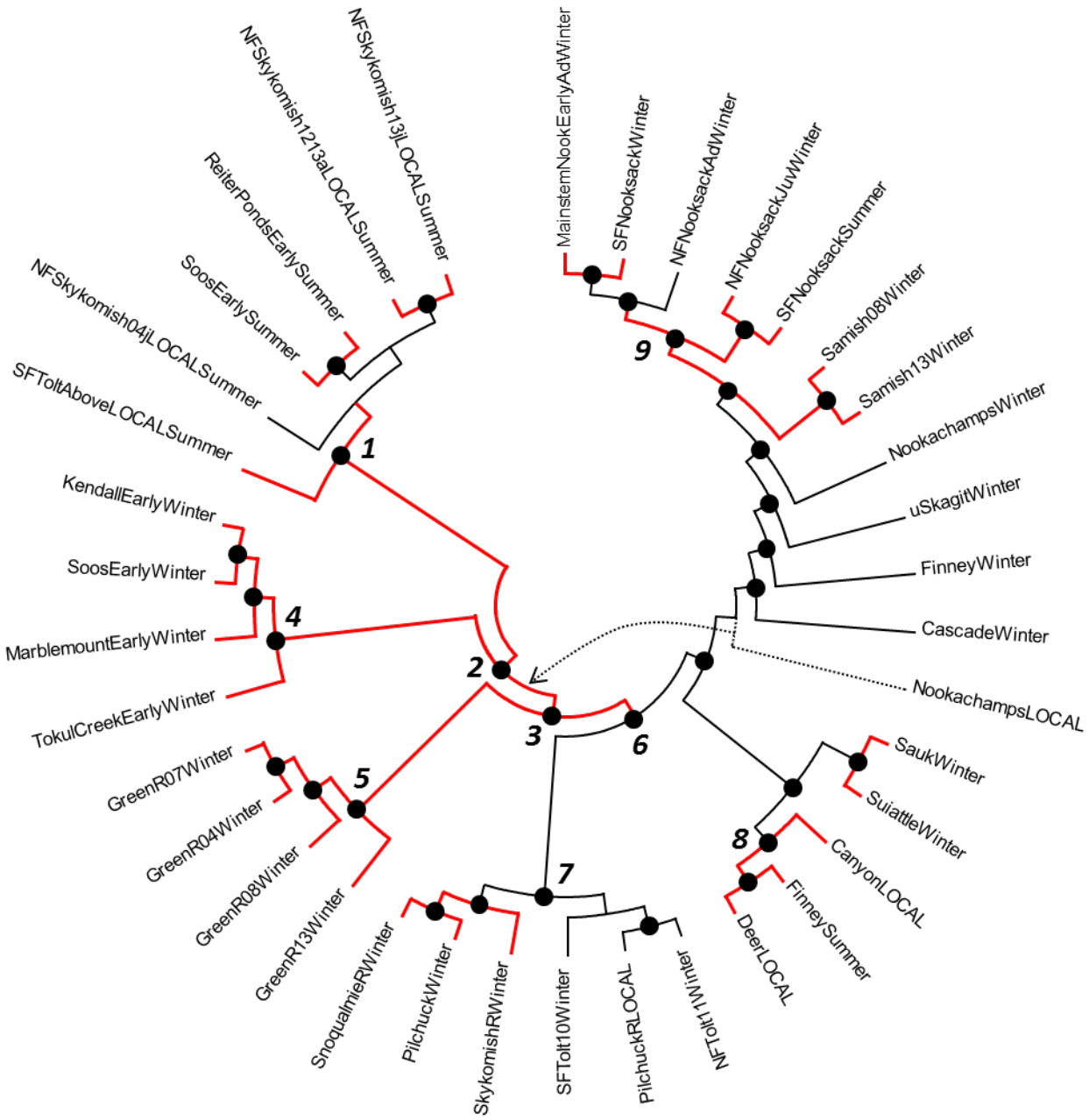

Figure 7. Neighbor-joining tree constructed from the pairwise Nei distance matrix of the taxa in the dataset described in the "Identity of Local Assignments" section in the Methods. The black enclosed circles on specific nodes signify nodes present in the bootstrap consensus tree (with one exception, see below), and red lines indicate branches extending from nodes that occur in greater than 45% of the 1000 iterations of the consensus tree. The numbered nodes show (1) early summer hatchery ("Skamania" or lower Columbia River) lineage (86% of bootstrap trees), (2) Puget Sound lineage (70%), (3) Puget Sound wild steelhead lineage (88%), (4) early winter ("Chambers") hatchery lineage (100%), (5) Green River wild lineage (49%), (6) Whidbey Basin, plus Nooksack group, including a divided polyphyletic and paraphyletic wild Skagit River "group" (21%), (7) Snohomish River wild lineage (17%), (8) wild summer lineage (perhaps Stillaguamish, depending on the ancestry of the FinneySummer fish) (51%), and (9) Nooksack River wild lineage (99%). In the consensus tree, the NookachampsLOCAL taxon appears where indicated by the dotted arrow (between nodes 2 and 3). The tree is rooted along the branch between nodes 1 and 2

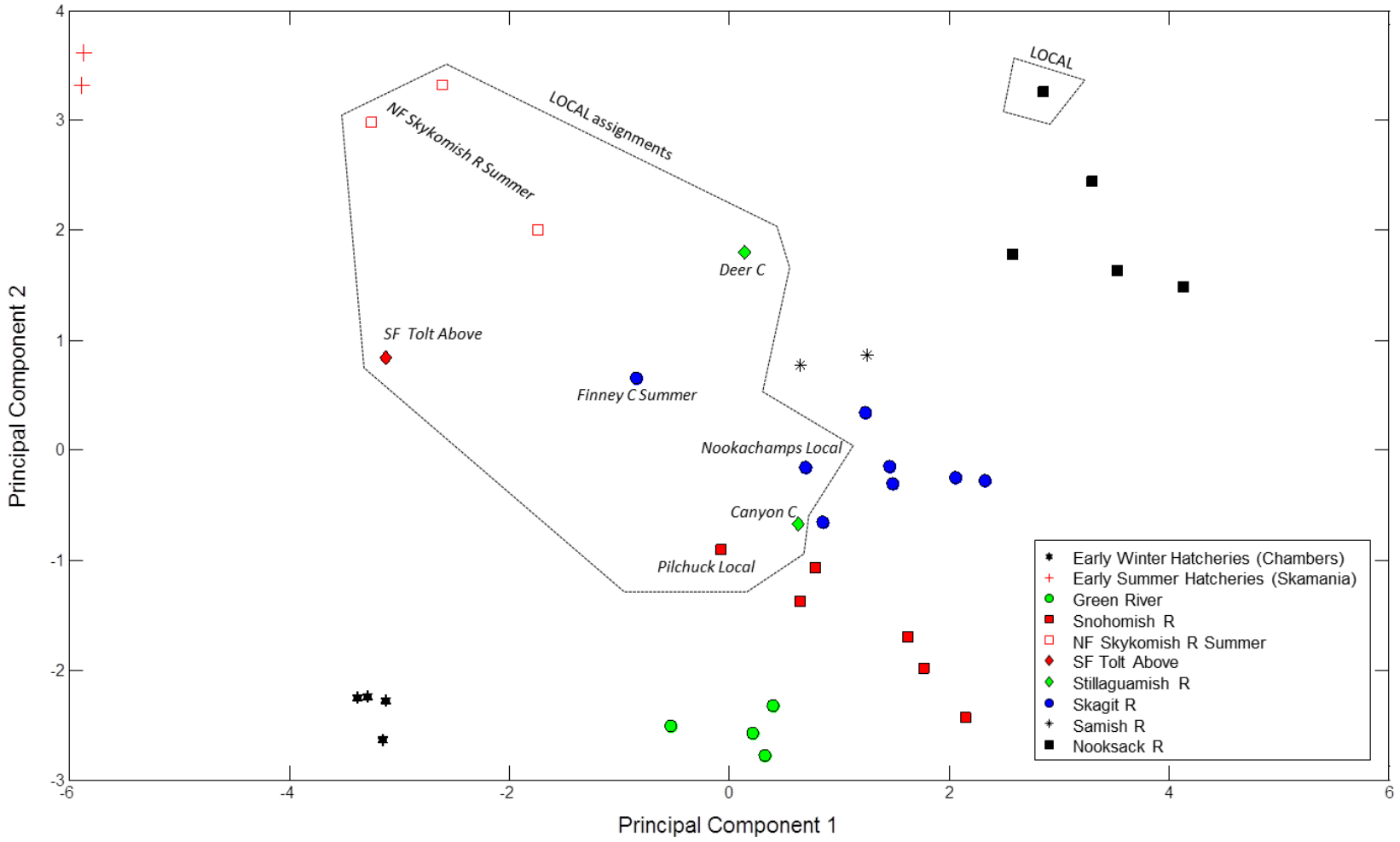

Figure 8. Principal component analysis (PCA) of the 37 taxa shown in Figure 7. Principal components 1 and 2 were the only significant (i.e., stable) components. The two polygons enclose all taxa designated by the *Structure* analyses as being "Local" (see text). PCA was conducted using individual allele frequencies, but plot shows only the centroids for each of the 37 taxa.

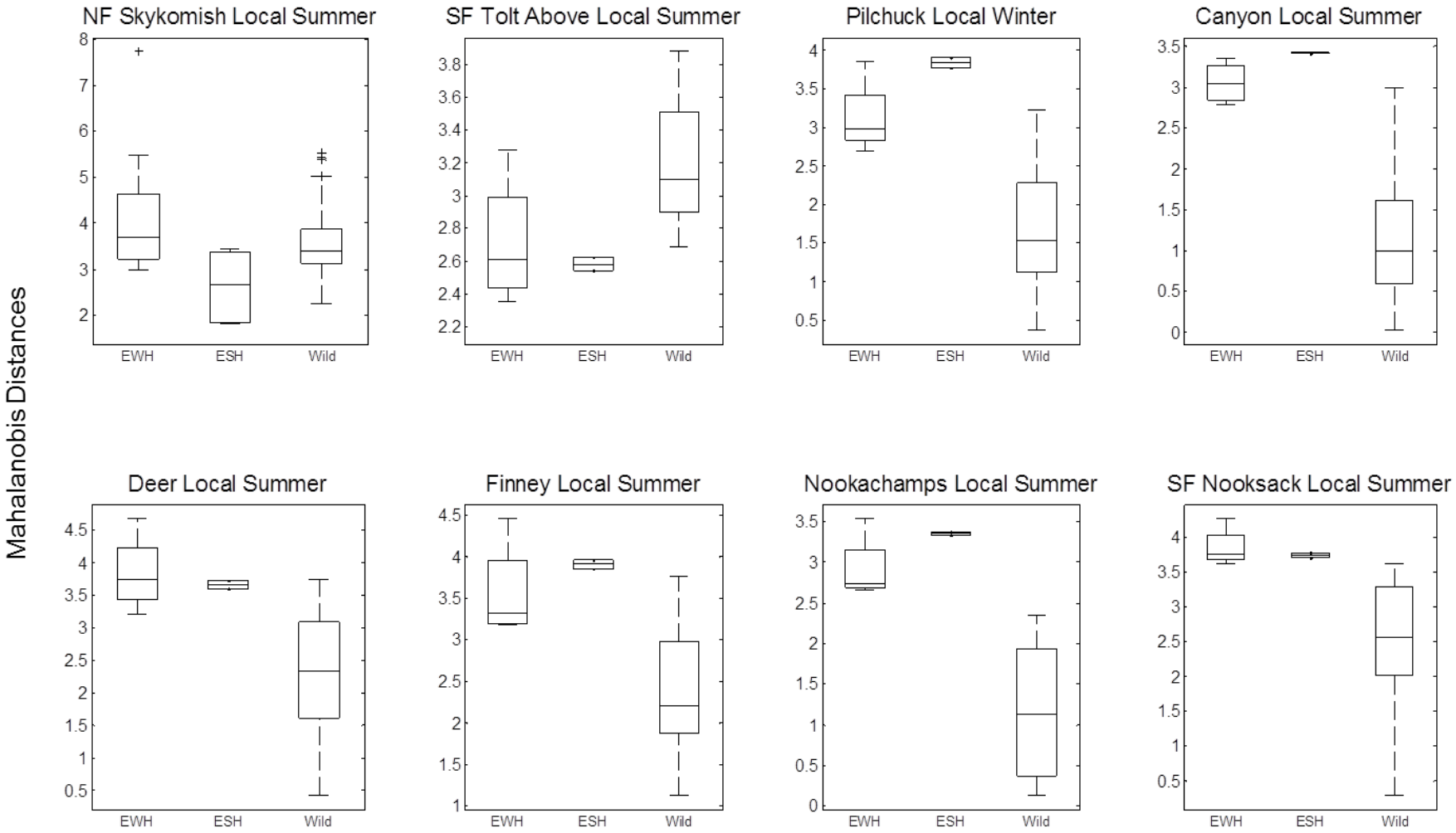

Figure 9. Box plots for pairwise Mahalanobis distances between all local taxa in Figure 8 (three NF Skykomish taxa were combined here), and the aggregate of all early winter hatchery (left box plot in each square), early summer hatchery (center box plot in each square), and non-local wild (right box plot in each square) taxa. Horizontal line in each box corresponds to the median value, lower and upper bounds of the box are the first and third quartile, respectively, the "whisker" tips cover approximately 99% of the data, if the data were normally distributed, and the "+" are outliers.

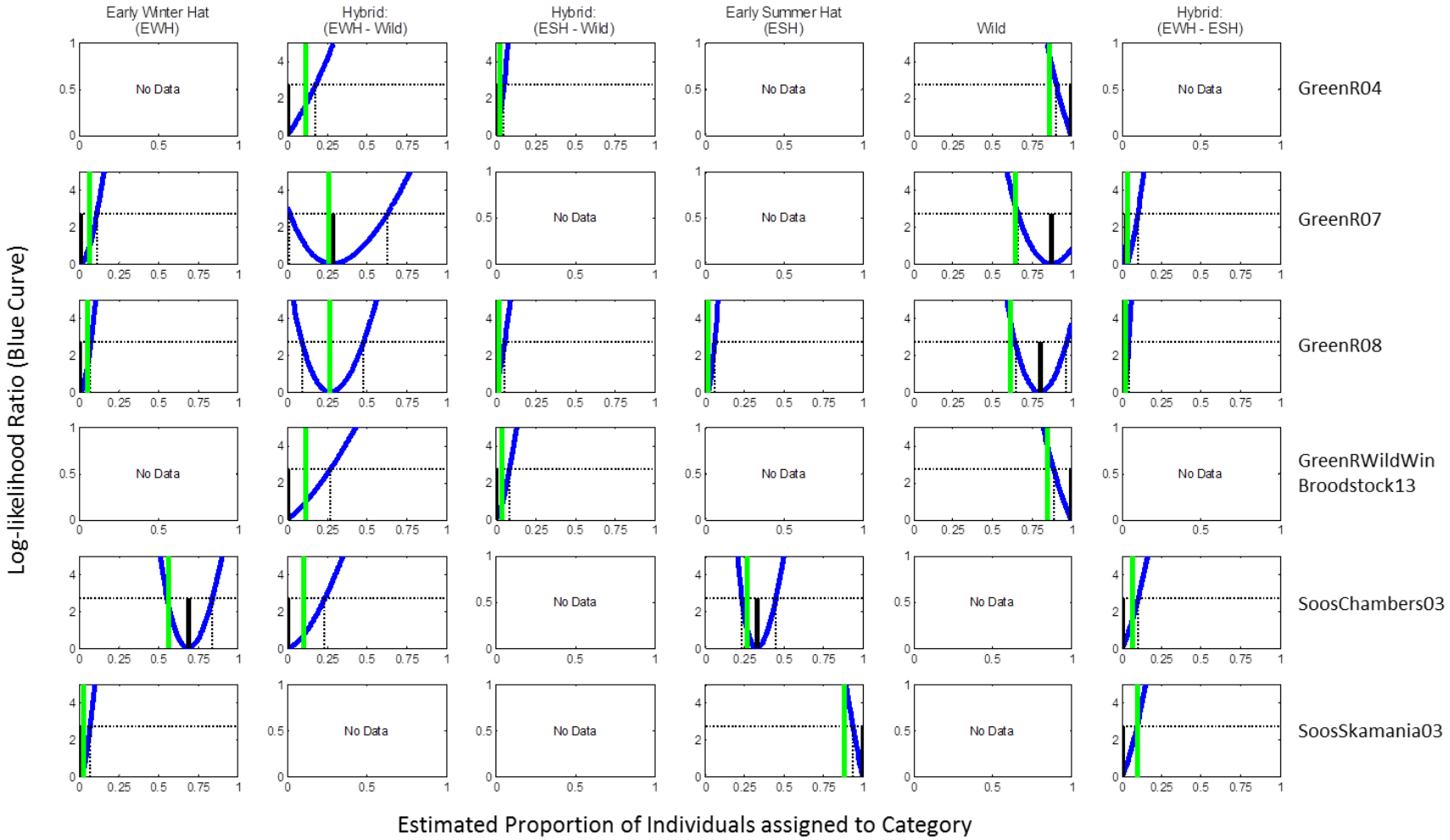

Figure 10a. Log-likelihood ratios (blue curve) for the estimated proportion of individuals assigned to each category (columns) for each Operational Unit (OU) within the Green River basin (rows). Black vertical line indicates the maximum likelihood estimate for the proportion, based on the likelihood function (Fig. 5, Steps 10-11) and the original *Structure* proportion (green vertical line). The horizontal distance between the black and green vertical lines shows the degree to which the *Structure* proportions were adjusted. Horizontal dotted line is the chi-square critical value, and the vertical dotted lines bound the 90% confidence interval for the maximum likelihood estimate (Fig. 5, Steps 12-13). If the black vertical line is not visible, the maximum likelihood estimate is either 0 or 1, depending on the position of the log-likelihood ratio curves (to the left or right, respectively). No Data indicates that there was no *Structure* proportion (i.e., proportion = 0) for that category. See Table 1b for description of OUs.

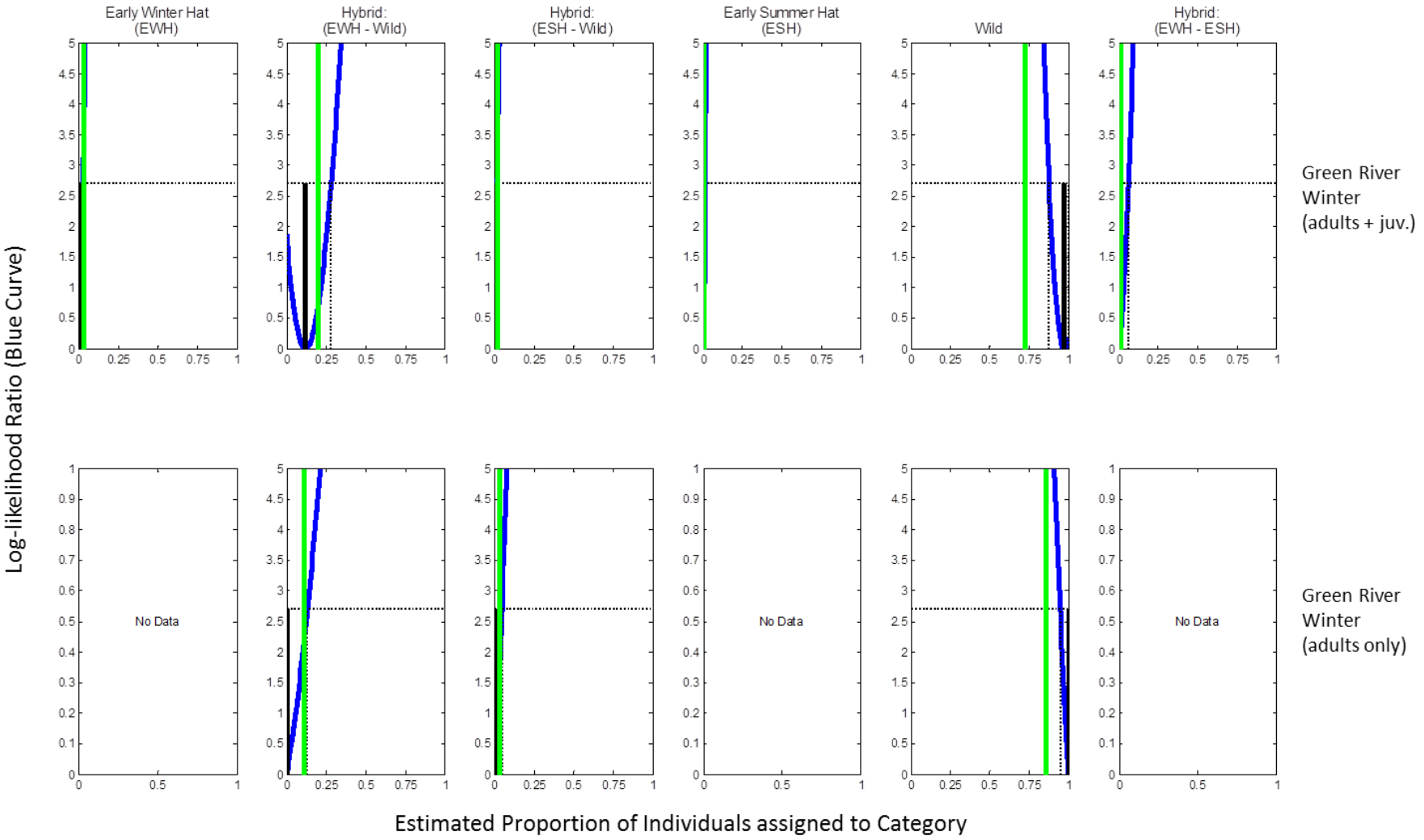

Figure 10b. Same as Figure 10a, except the rows correspond to the Green River Winter DIP including adults and juveniles (above), and adults only (below).

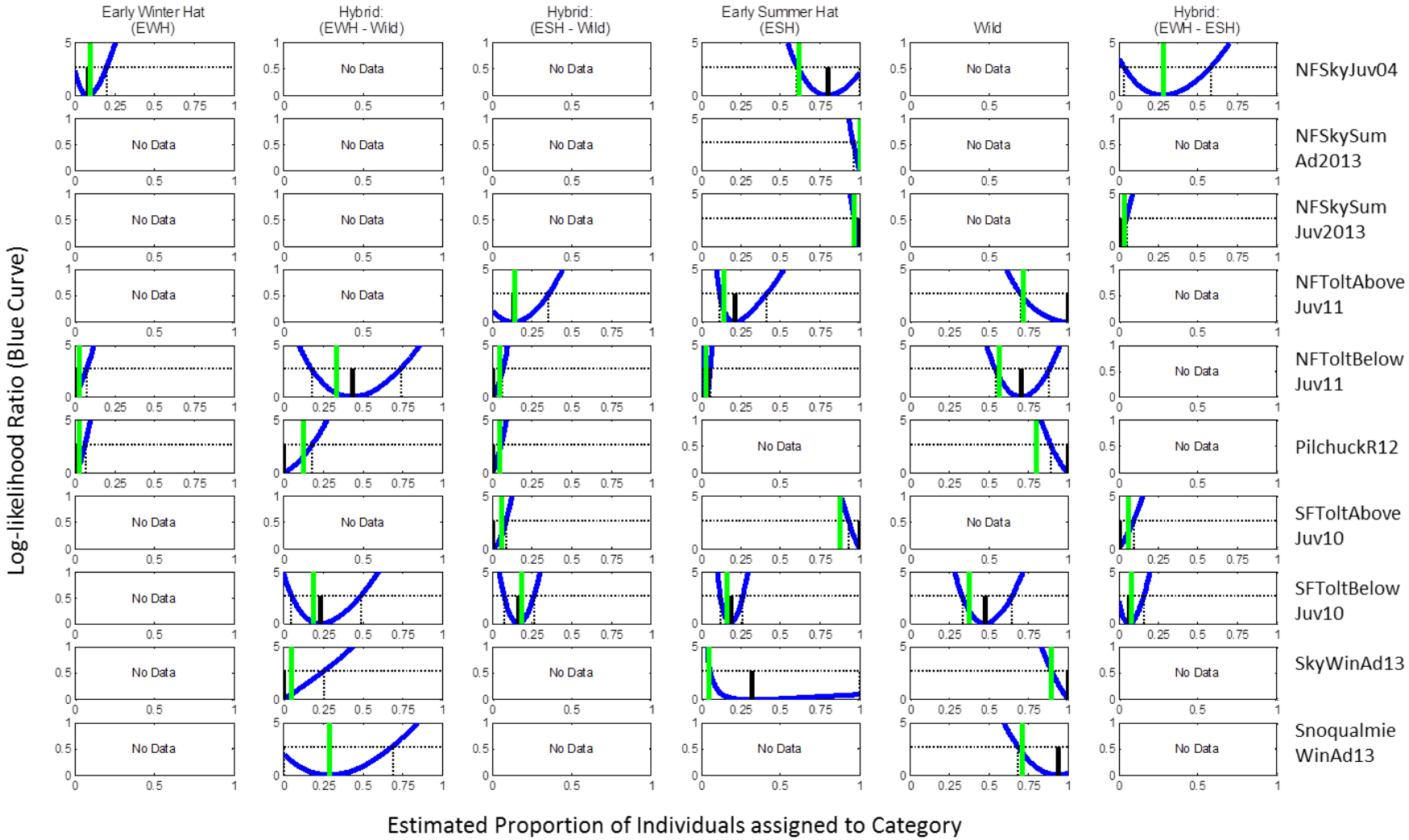

Figure 11a. Log-likelihood ratios (blue curve) for the estimated proportion of individuals assigned to each category (columns) for each Operational Unit (OU) within the Snohomish River basin (rows). Black vertical line indicates the maximum likelihood estimate for the proportion, based on the likelihood function (Fig. 5, Steps 10-11) and the original *Structure* proportion (green vertical line). The horizontal distance between the black and green vertical lines shows the degree to which the *Structure* proportions were adjusted. Horizontal dotted line is the chi-square critical value, and the vertical dotted lines bound the 90% confidence interval for the maximum likelihood estimate (Fig. 5, Steps 12-13). If the black vertical line is not visible, the maximum likelihood estimate is either 0 or 1, depending on the position of the log-likelihood ratio curves (to the left or right, respectively). No Data indicates that there was no *Structure* proportion (i.e., proportion = 0) for that category. See Table 1b for description of OUs.

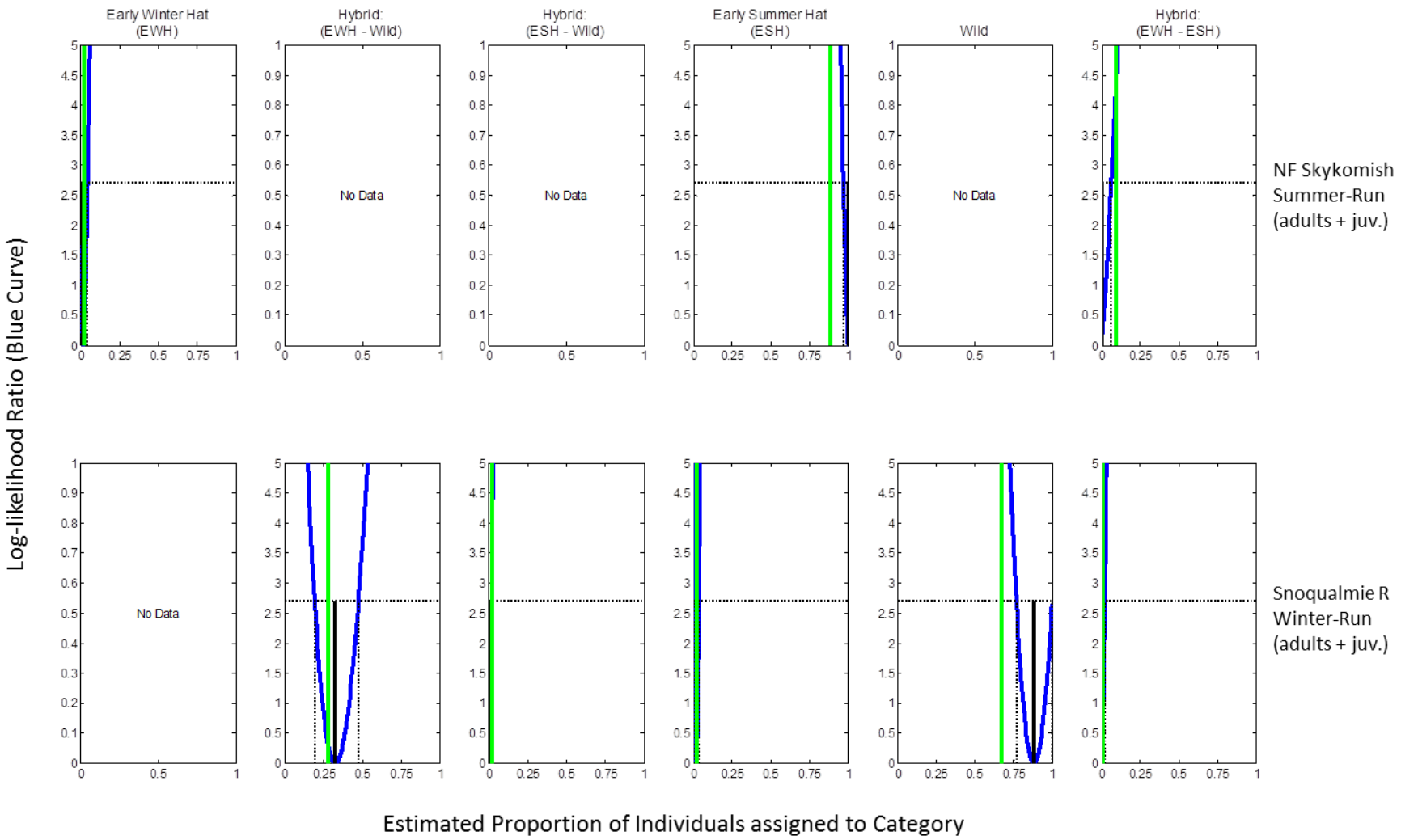

Figure 11b. Same as Figure 11a, except the rows correspond to the North Fork Skykomish Summer-Run (upper) and the Snoqualmie R Winter-Run (lower) DIPs, including adults and juveniles. The adult only versions of these analyses are monotypic and shown in Figure 11a as NFSkySumAd2013, and SnoqualmieWinAd13, respectively. The remaining three DIPs in the Snohomish, Pilchuck R Winter-Run (adults only), Snohomish/Skykomish R Winter-Run (adults only), and Tolt River Summer-Run (juveniles only), are also monotypic and shown in Figure 11a as PilchuckR12, AkyWinAd13, and SFToltAboveJuv10, respectively.

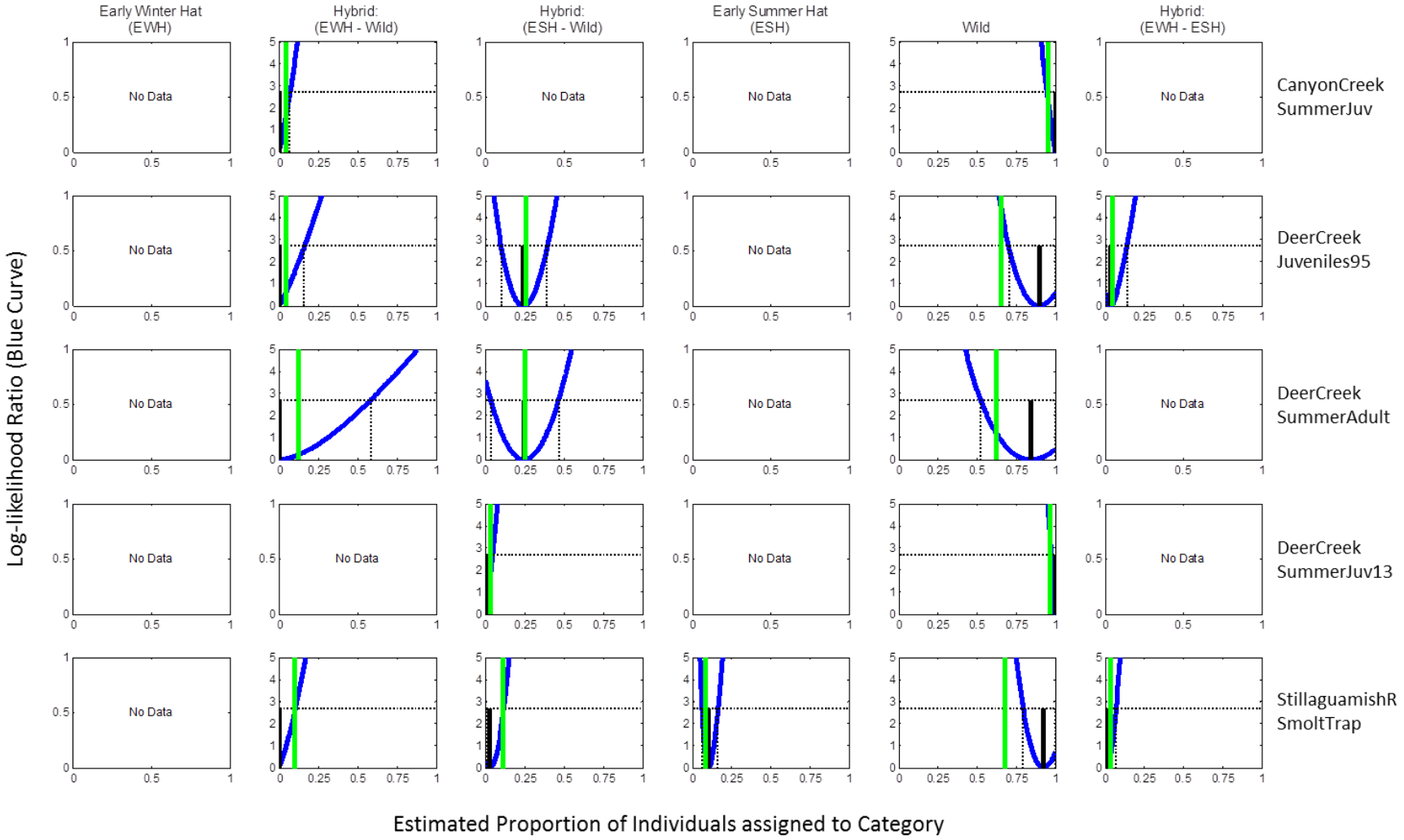

Figure 12a. Log-likelihood ratios (blue curve) for the estimated proportion of individuals assigned to each category (columns) for each Operational Unit (OU) within the Stillaguamish River basin (rows). Black vertical line indicates the maximum likelihood estimate for the proportion, based on the likelihood function (Fig. 5, Steps 10-11) and the original *Structure* proportion (green vertical line). The horizontal distance between the black and green vertical lines shows the degree to which the *Structure* proportions were adjusted. Horizontal dotted line is the chi-square critical value, and the vertical dotted lines bound the 90% confidence interval for the maximum likelihood estimate (Fig. 5, Steps 12-13). If the black vertical line is not visible, the maximum likelihood estimate is either 0 or 1, depending on the position of the log-likelihood ratio curves (to the left or right, respectively). No Data indicates that there was no *Structure* proportion (i.e., proportion = 0) for that category. See Table 1b for description of OUs.

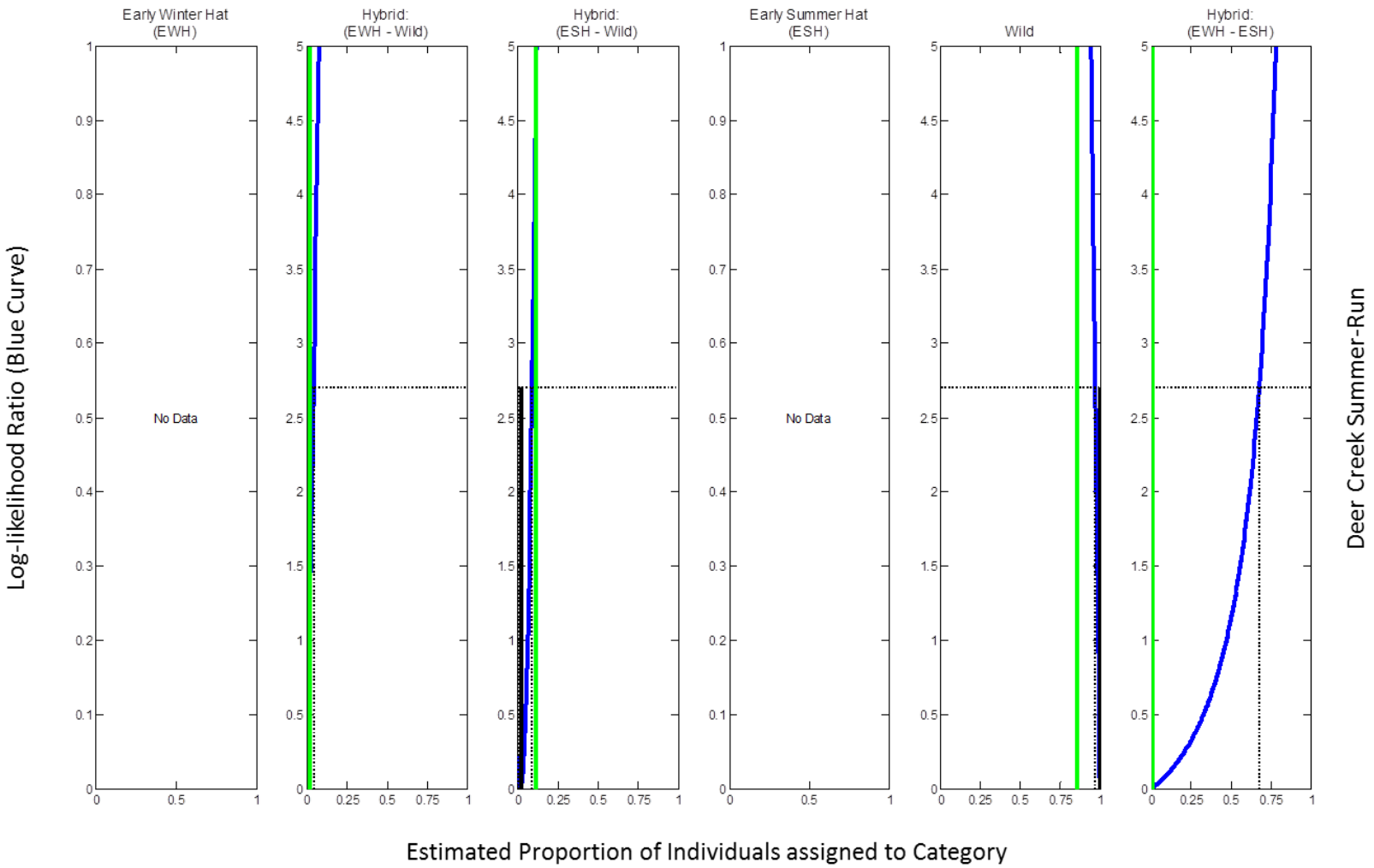

Figure 12b. Same as Figure 12a, except the row here correspond to the Deer Creek DIP, including adults and juveniles. The adult only version of this DIP is monotypic and shown in Figure 12a as DeerCreekSummerAdult. Canyon Creek Summer-Run (juveniles only) is the other Stillaguamish DIP analyzed here, and is shown in Figure 12a as CanyonCreekSummerJuv.

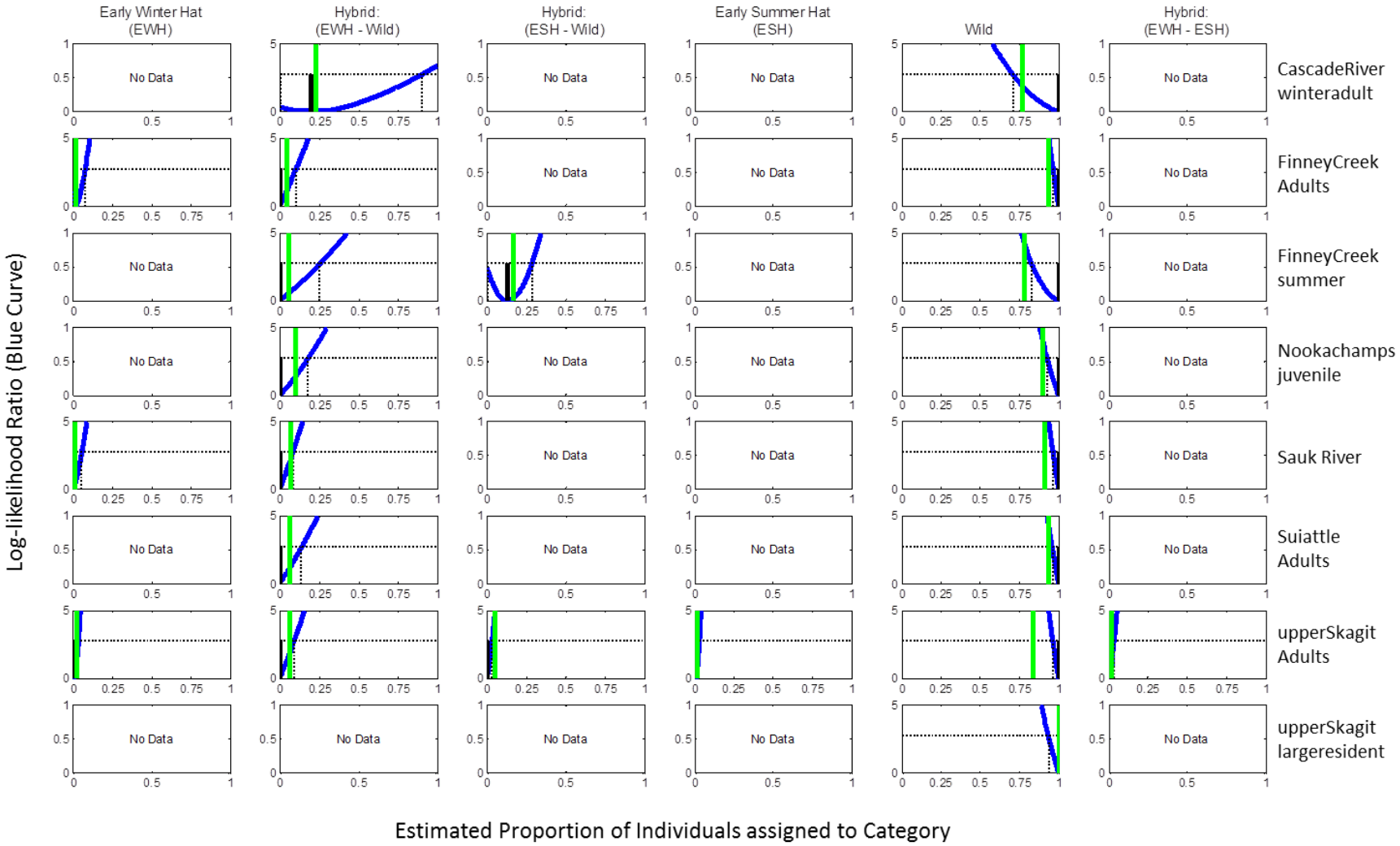

Figure 13a. Log-likelihood ratios (blue curve) for the estimated proportion of individuals assigned to each category (columns) for each Operational Unit (OU) within the Skagit River basin (rows). Black vertical line indicates the maximum likelihood estimate for the proportion, based on the likelihood function (Fig. 5, Steps 10-11) and the original *Structure* proportion (green vertical line). The horizontal distance between the black and green vertical lines shows the degree to which the *Structure* proportions were adjusted. Horizontal dotted line is the chi-square critical value, and the vertical dotted lines bound the 90% confidence interval for the maximum likelihood estimate (Fig. 5, Steps 12-13). If the black vertical line is not visible, the maximum likelihood estimate is either 0 or 1, depending on the position of the log-likelihood ratio curves (to the left or right, respectively). No Data indicates that there was no *Structure* proportion (i.e., proportion = 0) for that category. See Table 1b for description of OUs.

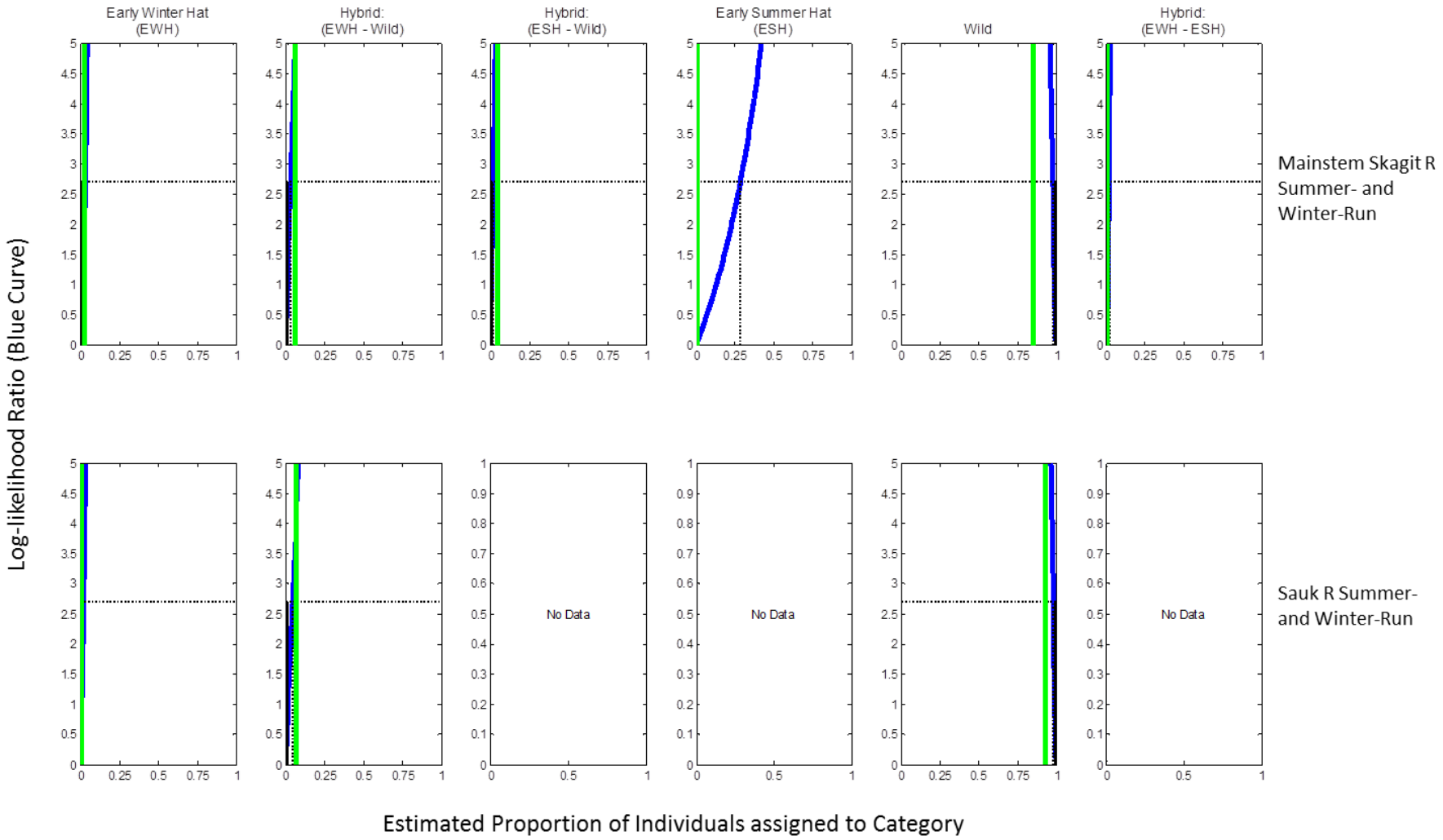

Figure 13b. Same as Figure 13a, except the rows correspond to the Mainstem Skagit R Summer- and Winter-Run (upper) and the Sauk R Summer- and Winter-Run (lower) DIPs, including adults only. Nookachamps Creek Winter-Run (juveniles only) is the other Skagit DIP analyzed here, and is shown in Figure 13a as NookachampsJuvenile.

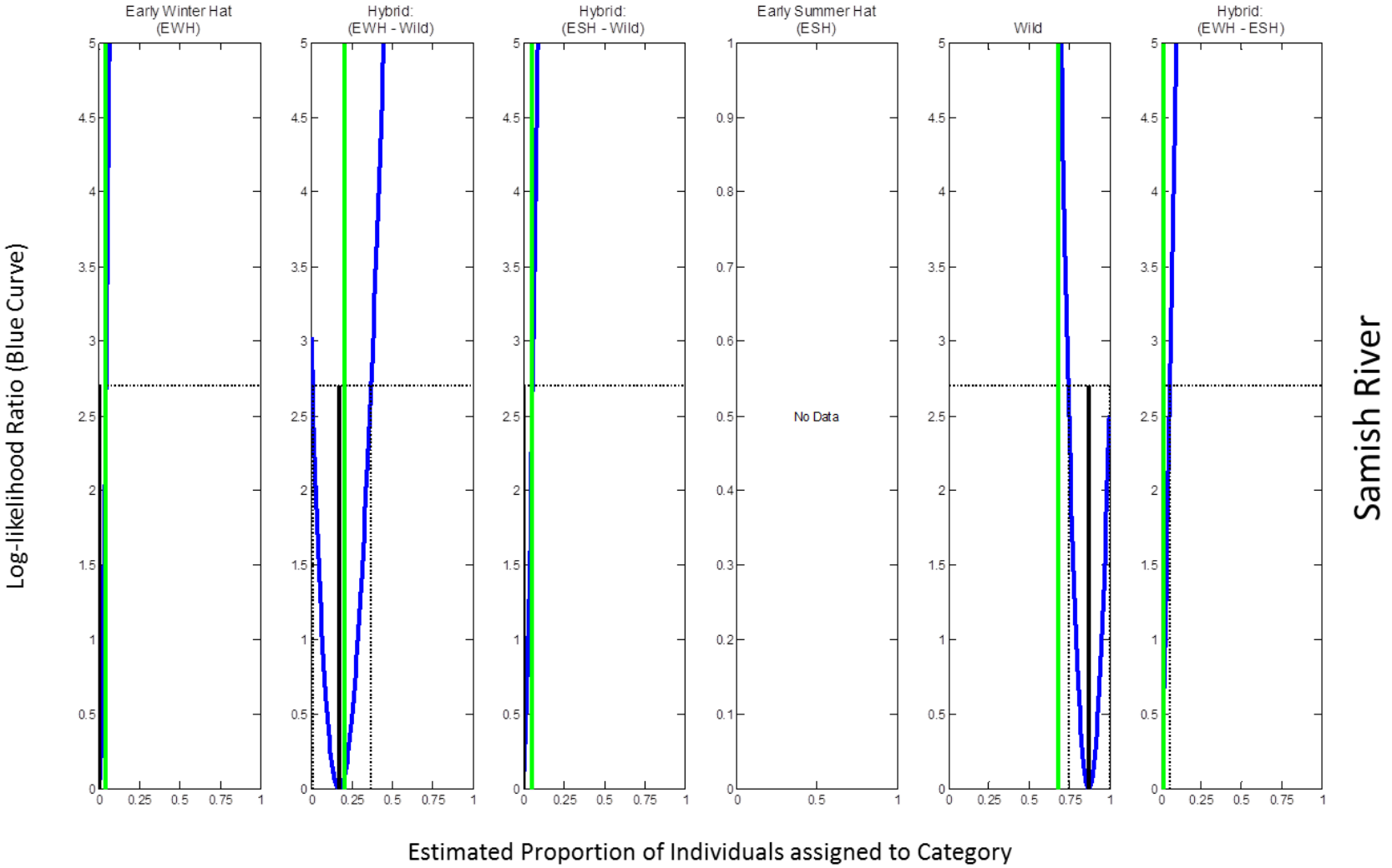

Figure 14. Log-likelihood ratios (blue curve) for the estimated proportion of individuals assigned to each category (columns) for the one Operational Unit (OU) and DIP (Samish River Winter Run) within the Samish River. Black vertical line indicates the maximum likelihood estimate for the proportion, based on the likelihood function (Fig. 5, Steps 10-11) and the original *Structure* proportion (green vertical line). The horizontal distance between the black and green vertical lines shows the degree to which the *Structure* proportions were adjusted. Horizontal dotted line is the chi-square critical value, and the vertical dotted lines bound the 90% confidence interval for the maximum likelihood estimate (Fig. 5, Steps 12-13). If the black vertical line is not visible, the maximum likelihood estimate is either 0 or 1, depending on the position of the log-likelihood ratio curves (to the left or right, respectively). No Data indicates that there was no *Structure* proportion (i.e., proportion = 0) for that category.

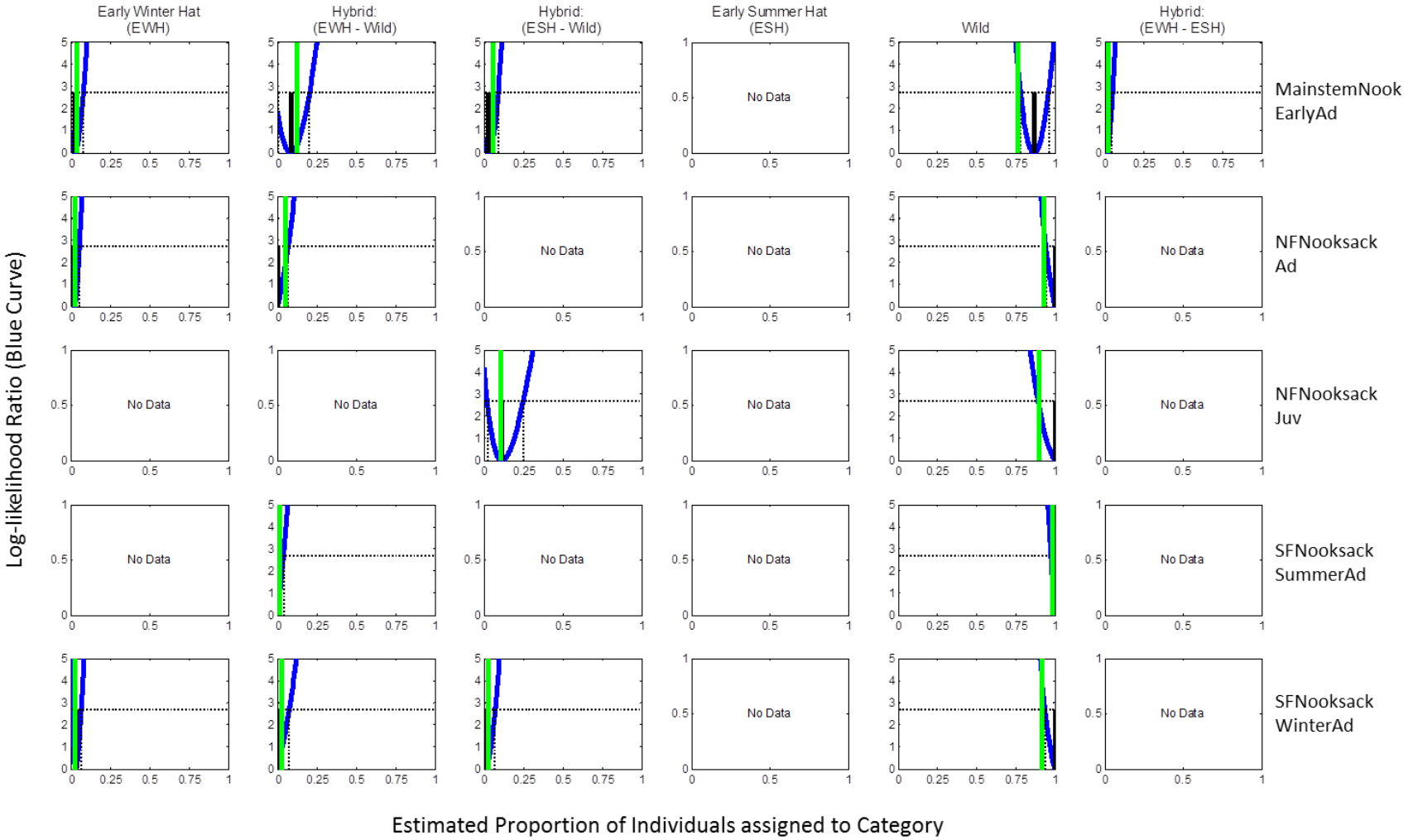

Figure 15a. Log-likelihood ratios (blue curve) for the estimated proportion of individuals assigned to each category (columns) for each Operational Unit (OU) within the Nooksack River basin (rows). Black vertical line indicates the maximum likelihood estimate for the proportion, based on the likelihood function (Fig. 5, Steps 10-11) and the original *Structure* proportion (green vertical line). The horizontal distance between the black and green vertical lines shows the degree to which the *Structure* proportions were adjusted. Horizontal dotted line is the chi-square critical value, and the vertical dotted lines bound the 90% confidence interval for the maximum likelihood estimate (Fig. 5, Steps 12-13). If the black vertical line is not visible, the maximum likelihood estimate is either 0 or 1, depending on the position of the log-likelihood ratio curves (to the left or right, respectively). No Data indicates that there was no *Structure* proportion (i.e., proportion = 0) for that category. See Table 1b for description of OUs.

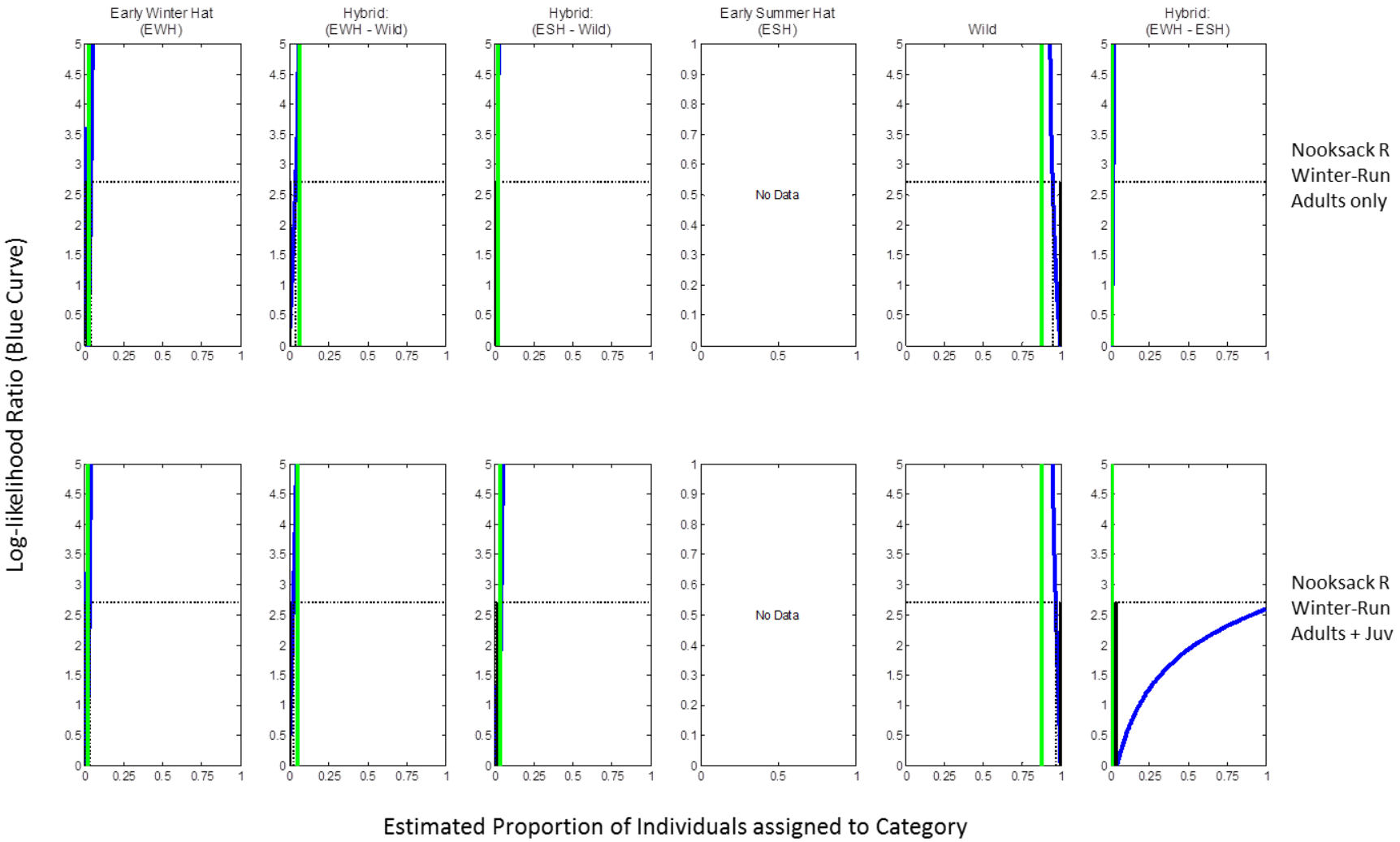

Figure 15b. Same as Figure 15a, except the rows correspond to the Nooksack R Winter-Run including adults only (above), and adults and juveniles (below).. South Fork Nooksack R Summer-Run (adults only) is the other Nooksack DIP analyzed here, and is shown in Figure 15a as SFNooksackSummerAd.

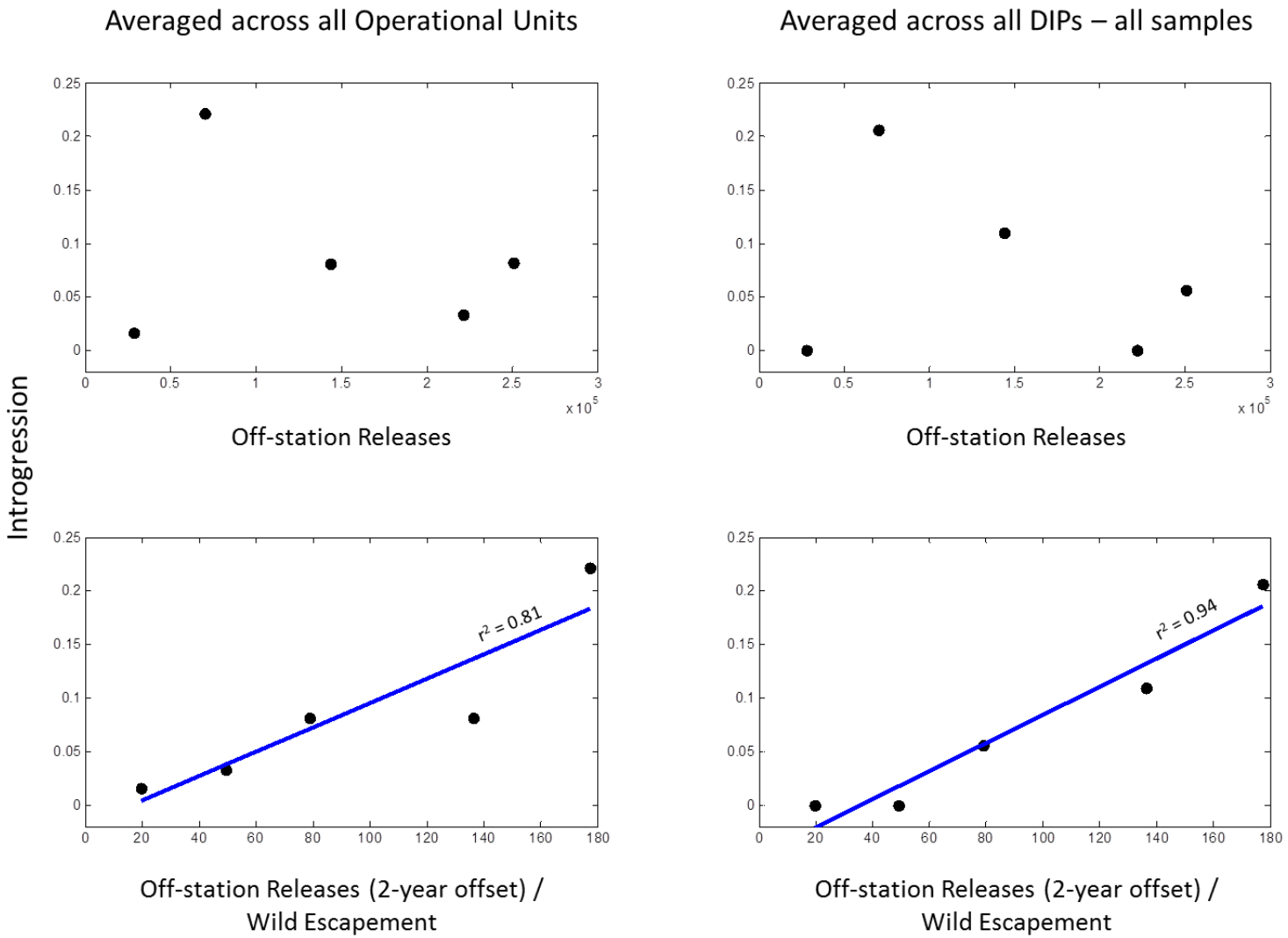

Figure 16. Introgression, averaged across all operation units (left two plots) or DIPs (right two plots), as a function of average number of off-station releases (top two plots) or ratio between average number of off-station releases and wild escapement (bottom two plots). Off-station releases were offset by two years, corresponding to the year when the released smolts would return as adults. Where available, releases included years 2000 – 2010 (corresponding to return years 2002 – 2012), and escapement included years 2002 – 2012. Reading from high to low introgression for the operational units, the points correspond to: Snohomish River summer – run (Reiter Ponds releases), Snohomish River winter – run (Tokul Creek releases), Green River winter-run (Soos Creek - EWH releases), Skagit River winter – run (Marblemount releases), and Nooksack River total – run (Kendall Creek releases). Reading from high to low introgression for the DIPs: Snohomish River summer – run, Green River winter-run, Snohomish River winter – run, and Skagit River winter – run and Nooksack River total – run both with zero introgression.