1

2   DR. JON E. HESS (Orcid ID : 0000-0002-3643-202X)

3   DR. THOMAS A. DELOMAS (Orcid ID : 0000-0001-5154-759X)

4

5

6   Article type   : Article

7

8

9   Robust recolonization of Pacific lamprey following dam removals

10   Jon E. Hess[1]*, Rebecca L. Paradis[2], Mary L. Moser[3], Laurie A. Weitkamp[4], Thomas A. Delomas[5],

11   and Shawn R. Narum[6]

12

13

14   Full postal addresses:

15   1. Columbia River Inter-Tribal Fish Commission; 700 NE Multnomah Street; Suite 1200;

16   Portland, OR 97232

17   2. Lower Elwha Klallam Tribe; 760 Stratton Road; Port Angeles, WA 98363

18   3. Northwest Fisheries Science Center; NOAA Fisheries; 2725 Montlake Blvd E.; Seattle

19   WA 98112

20   4. Northwest Fisheries Science Center; NOAA Fisheries; 2032 Marine Science Dr.;

21   Newport, OR 97365

22   5. Pacific States Marine Fisheries Commission / Idaho Department of Fish and Game, Eagle

23   Fish Genetics Laboratory, 1800 Trout Rd. Eagle, ID, 83616

24   6. Columbia River Inter-Tribal Fish Commission; 3059-F National Fish Hatchery Rd;

25   Hagerman, ID 83332

26   *corresponding author, to whom proofs will be sent:

27   Jon E. Hess

28   Columbia River Inter-Tribal Fish Commission

29   700 NE Multnomah Street; Suite 1200; Portland, OR 97232

30   Fax: (503)731-1307

31   e-mail: hesj@critfc.org

32

33   RUNNING TITLE: Pacific lamprey recolonization

34   Abstract

35   Removal of two dams in the Elwha River basin started one of the largest river restoration

36   projects ever attempted in the Pacific Northwest. These dams had eliminated Pacific lamprey

37   populations upstream. After the dam removals, larval production increased in the upper

38   watershed, but the sources and numbers of new adult spawners were unknown. We applied

39   genetic stock identification (GSI), parentage (PA), and sibship (SA) assignment methods to: 1)

40   determine the origins of Pacific lamprey larvae and juveniles, 2) quantify the increase in

41   numbers of successful Elwha River spawners (i.e., "effective spawners", N_b) and assess whether

42   the current numbers of spawners have reached levels equivalent to those of neighboring

43   undammed basins, and 3) determine the relative productivity of streams within the Elwha River

44   and how overall productivity originating from this system may be distributed across the broader

45   surrounding region. We utilized a highly accurate set of 263 single nucleotide polymorphism

46   (SNP) loci to perform PA and SA (>99% accuracy rate); and an additional set of 28 SNPs for

47   GSI. Our results showed that a single stream (Indian Creek) was the source of 41% of larval and

48   juvenile production in the Elwha River. Our N_b estimates for the Elwha River indicated a 12-

49   fold increase in N_b in the three years following dam removal, with recent N_b estimates matching

50   those of neighboring Olympic Peninsula basins. These results indicate rapid recolonization

51   potential for this highly dispersive species and high productivity within the Elwha River suggests

52   that restoring passage to adequate habitat is a highly effective approach to re-establish

53   populations of Pacific lamprey in coastal systems.

54   Introduction

55   Pacific lamprey have suffered declines across the species' range due to habitat loss,

56   passage impediments, and even active poisoning. Multifaceted strategies have been employed to

57   re-establish self-sustaining, harvestable abundance of adults, larvae, and juveniles. The greatest

58   level of human intervention may be required in the upper reaches of large river basins. For

59   example, the Columbia River Treaty Tribes have employed high levels of sustained translocation

60   (i.e., human-mediated transfers of fish to areas upstream from hydropower dams) and developed

61 effective artificial propagation. In smaller coastal systems, the species may need just an

62 opportunity to access appropriate habitat to naturally re-establish because these areas have fewer

63 passage impediments and greater proximity to oceanic sources of Pacific lamprey spawners. For

64 small coastal rivers with dams that prevent passage of Pacific lamprey, dam removal could be the

65 only requisite for restoration of robust and self-sustaining Pacific lamprey abundance via natural

66 recolonization (Reid & Goodman 2020; Moser et al. In Press).

67 Pacific lamprey in the Elwha River drainage present an ideal case study for examining

68 the characteristics of anadromous recolonization of a coastal river basin following dam removal.

69 Removal of two large, obsolete dams started one of the largest river restoration projects ever

70 attempted in the Pacific Northwest. These dams had eliminated Pacific lamprey populations

71 upstream, but adults rapidly recolonized the upper watershed after dam removal (Moser and

72 Paradis 2017). While natural recolonization was underway, adult lamprey from neighboring

73 watersheds were collected, genetic sampled, radio-tagged and translocated to the lower Elwha

74 River to document migratory behavior. These collections allowed evaluation of the effectiveness

75 of translocating fish during active natural recolonization efforts.

76 While translocation of adult Pacific lamprey can aid population recovery (Close et al.

77 2009; Ward et al. 2012), natural recolonization following barrier removal can also occur rapidly

78 (e.g., Hogg et al. 2013; Hess et al. 2015; Pereira et al. 2017; Jolley et al. 2018). Lampreys

79 apparently are not philopatric and use pheromonal cues to find watersheds for spawning

80 (reviewed in Moser et al. 2015). The relative rates of population recovery following natural

81 recolonization compared to production resulting from translocated adults likely depends on the

82 system and sources of lamprey (Reid and Goodman 2020).

83 Immediately after dam removal was completed in March 2012, Pacific lamprey were

84 noted upstream from the former Elwha Dam site (Moser & Paradis 2017), however, the exact

85 numbers and sources of effective spawners responsible for this production were unknown. The

86 larvae in one stream (Indian Creek) were observed to increase by 14X in the five years following

87 dam removal (from 132 larvae in 2014 to 1,805 larvae in 2016; Moser & Paradis 2017), but it

88 was unknown whether effective spawners increased at the same rate. Given the high fecundity

89 of Pacific lamprey (Kan 1975), it is possible that the increase in spawners could be much lower

90 than the observed increase in larvae, indicating a less robust recovery.

91   Second, it was unclear what the relative productivity was across the streams within the

92   Elwha River basin because all streams could not be monitored intensively and consistently. One

93   stream, Indian Creek, appears to have supported relatively high increases of juvenile salmonids

94   (i.e., Chinook Salmon *Oncorhynchus tshawytscha*, Coho Salmon *O. kisutch*, and Steelhead *O.*

95   *mykiss*; McHenry et al. 2020) compared to other streams of the Elwha River basin. A similar

96   pattern may occur in Pacific lamprey recruitment.

97   Finally, the relative contribution of this new productivity in the Elwha basin could benefit

98   the entire region given the lack of homing in this species. Understanding sources and sinks of

99   production is a key management question for all native anadromous lampreys (Lucas et al. In

100   Press). Until recently, understanding population dynamics of anadromous lampreys was

101   rudimentary at best. Yet such information is crucial for adaptive management of these imperiled

102   species. With the development of recent genetic methods for Pacific lamprey, like genetic stock

103   identification (GSI) and parentage (PA) and sibship assignment (SA), this information has

104   become available.

105   The three different genetic methods (GSI, PA, and SA) provide ways to address similar

106   objectives (e.g. identify natal origins of individuals), but they differ in precision and sampling

107   requirements. GSI identifies individuals to their genetic stock-of-origin and can be an effective

108   way to perform mixed stock analysis for management (Shaklee et al. 1999). The precision of

109   GSI depends on the spatial scale in which stock level differentiation is observed, which can be at

110   small scales within the basin level for some species (e.g. steelhead in the Snake River basin,

111   Nielsen et al. 2009; Campbell et al. 2012), but at larger regional spatial scales for others (e.g.

112   Eulachon, *Thaleichthys pacificus*; Candy et al. 2015). PA and SA can be used to precisely

113   identify individuals to the spawning-site-of-origin of the sampled parents (Steele et al. 2019) or

114   natal-site-of-origin of sampled siblings in the baseline (Jones and Wang 2010), respectively. In

115   addition, SA can be used to estimate effective numbers of spawners (N_b), a potential proxy for

116   adult abundance (Ackerman et al. 2017). Larval and juvenile full-sibling baselines can be

117   effective to estimate the number of effective spawners that contributed to each collection of

118   young life stages (e.g., Hess et al. 2015; Whitlock et al. 2017). Here we define "larvae" as

119   immature, filter-feeding lamprey without eyes and "juveniles" as fully transformed individuals

120   with eyes that are ready to feed (Clemens 2019).

121   We used genetic monitoring of the Elwha River Pacific lamprey to address the following

122   three objectives: 1) determine the origins of larvae and juveniles of Pacific lamprey collected in

123   the Elwha River and thereby attribute production to translocation versus natural recolonization,

124   2) quantify the increase in numbers of effective spawners in the Elwha River and determine

125   whether current numbers of effective spawners have reached levels observed in neighboring

126   undammed basins, and 3) determine the relative productivity of streams within the Elwha River

127   basin and how overall productivity originating from the Elwha River is distributed across the

128   broader surrounding region.

129   We used a highly accurate (>99% accuracy rate) set of 263 single nucleotide

130   polymorphism (SNP) loci to perform single PA and SA, and an additional 28 SNPs to perform

131   GSI. We expected the following answers to our questions: 1) Larval and juvenile production in

132   the Elwha River is primarily sourced from natural recolonization given the relatively modest

133   translocation effort (<50 adults), 2) increases in effective spawners is as rapid as the 14X

134   increase in larvae during the five years following dam removal; however, given the relatively

135   recent recolonization in the Elwha River, numbers of effective spawners will be lower than in

136   undammed basins, and 3) within the Elwha River, Indian Creek would contribute most

137   production given its high productivity for salmonids, but the overall contribution of Pacific

138   lamprey from the Elwha River to the broader surrounding region would be relatively small given

139   its recent founding and small area relative to surrounding sources of Pacific lamprey.

140

141   **Methods**

142   *Sample collections*

143   The collections of Pacific lamprey included individuals from within the Elwha River

144   basin, neighboring basins of the Olympic Peninsula, ocean-phase individuals along the U.S.

145   West Coast, and adult returns to the Willamette River in the Columbia River (Fig. 1). In addition,

146   "Range-wide" reference collections that were previously genotyped were utilized for the GSI

147   baseline (Hess et al. 2013, labeled in Fig.1a).

148   The bulk of the samples were larvae, juveniles and adults collected from the Elwha River

149   (N=773). We referred to groups of these collections from the Elwha River as: "Elwha_EF" =

150   larval and juvenile Pacific lamprey that were electro-fished; "Elwha_ST" = larval and juvenile

151   Pacific lamprey that were collected via rotary screw trap; and "Elwha_pbt vol" = volitional adult

152 migrants used as candidate parents in the parentage based tagging baseline. One sample of

153 adults was translocated into the Elwha River from other basins (*N*=42, referred to as "Elwha_pbt

154 trans" in Table 1). All larvae and juveniles were collected via electrofishing or by rotary screw

155 trap; the latter method also recovered some post-spawned adults. Adults translocated into the

156 Elwha River from neighboring basins (i.e., Deep Creek, East Twin River, and Salt Creek) were

157 always released and radio-tagged the same day of capture.

158 Pacific lamprey samples were also collected from outside the Elwha River basin.

159 Electrofished samples from neighboring drainages (Deep Creek, East Twin River, West Twin

160 River, and Lyre River 2015-2018, N=355) are referred to as "OLY_EF" and were Olympic

161 Peninsula larvae and juveniles. One group of these genetic samples came from electro-fished

162 larvae in Salt Creek (#9,10) and were subsequently translocated in 2018 to Little River (site #36,

163 Fig. 1), a stream within the Elwha River basin that had no prior lamprey production. Genetic

164 samples were also taken from parasitic, ocean-phase individuals ("ocean", N=695) collected in

165 the 2017 Pacific hake (*Merluccius productus*) trawl fishery in the ocean off Washington, Oregon,

166 and northern California (Fig. 1). Samples from miscellaneous adults that were hand collected at

167 Willamette Falls (Columbia River drainage, Fig. 1) in 2014 and 2015 were used as a control

168 group to test assignment methods (N=1714, "MISC_Adult").

169

170 *Molecular analysis*

171 We used both ethanol and dry storage of fin tissues followed by extraction and

172 Polymerase Chain Reaction (PCR) for genotyping-by-sequencing techniques on an Illumina

173 sequencer. Single nucleotide polymorphism (SNP) genotypes were generated using the

174 Genotyping-in-Thousands by sequencing (GT-seq) custom amplicon method described by

175 Campbell et al. (2015). The 308 SNP loci in the panel were selected to be representative of

176 neutral and adaptive loci across the geographic range of Pacific lamprey, as described by Hess et

177 al. (2020). For this study, we used a subset of 295 of these previously published loci. For the

178 most part this subset was a reduction from the larger number of loci, however seven loci were

179 reanalyzed to target a different SNP than had been previously genotyped in the same amplicon

180 (Etr_480-67, Etr_930-35, Etr_2016-70, Etr_3037-68, Etr_6318-70, Etr_7166-73, and Etr_7974-

181 70). The final 295 SNP panel included 263 SNPs that were used for PA and SA applications (12

182 of these 263 SNPs are adaptive), four for species determination (Hess et al. 2015, Arakawa et al.

183   in review), and 28 additional SNPs for characterizing adaptive variation (defined by outlier tests;

184   Hess et al. 2013, 2020).

185

186   *Statistical analysis*

187   Using GENEPOP v.4.2, we performed linkage disequilibrium (LD) tests for all locus

188   pairs and characterized minor allele frequency (MAF). For the LD tests, the Range-wide

189   reference collections were pooled into the following four major groups to strike a balance

190   between maximizing samples sizes per test population and minimizing Wahlund effects (Hess et

191   al. 2013): North (A, B) , South (P, Q), Lower Columbia (G2, G3, H, I) and Interior Columbia

192   (J2, J3, K, L; Table S1; Hess et al. 2013). We used three software programs to conduct PA using

193   the 263 loci: parent pair assignments using SNPPIT v.1.0 (Anderson 2012), and single parent

194   assignments using COLONY v.2.0.6.5 (Jones and Wang 2010) and SEQUOIA v. 2.0.7 (Huisman

195   2017). These three programs were used in combination for parentage to support results through

196   concordance. SNPPIT parent pair assignments generally provide highest confidence but can

197   only be performed when both parents of a candidate offspring are present in the baseline. When

198   one parent is missing, single parent assignments must be relied upon to identify offspring from a

199   baseline of candidate parents and require greater power to match the same level of accuracy of

200   SNPPIT. Therefore, we compared any candidate parent assignments in COLONY with results

201   from SNPPIT parent pair assignments and SEQUOIA single parent assignments for concordance.

202   Further, we used new methods to perform simulations in SEQUOIA to measure the accuracy of

203   this approach given the information content (variation) in our markers and datasets. COLONY

204   was also used to reconstruct full-sibling families (Wang 2004) for each study collection and

205   estimate effective number of spawners (N_b, Wang 2009) using the 263 SNPs (Table 1).

206   Individuals were assigned as parent-offspring trios using a likelihood-based method

207   implemented in SNPPIT with the parameter –max-par-miss set to 262 (total SNP N -1). This

208   effectively allows all parents and offspring in the dataset to be compared regardless of missing

209   data (missing data were already minimized by excluding individuals missing >=10% genotypes).

210   We used an estimated SNP genotyping error rate of 1.0% or a per allele rate of 0.5%. The

211   SNPPIT assesses confidence of parentage assignments using a false discovery rate (FDR), and

212   we only accepted assignments with a very stringent FDR threshold of <0.5% (i.e. we expected

213   the fraction of offspring assigned to incorrect parents to be <1 in 200). A 0.5% per allele error

214   rate was chosen as a conservative value for this SNPPIT parameter because it is larger than the

215   observed per locus error rate of 0.2%.

216   In COLONY, we analyzed the larval and juvenile collections together as one (i.e.,

217   N=1815 for ELWHA_EF, ELWHA_ST, OLY_EF, and Ocean; Table 1) using the following

218   parameter settings: polygamous mating for males and females without inbreeding, full-likelihood,

219   medium length of run, no allele updating, and no sibship priors. These parameter settings were

220   chosen based on prior work that empirically demonstrated the polygamous system of mating and

221   confirmed the accuracy of the sibship reconstruction based on comparisons with known familial

222   relationships (Hess et al. 2015). We also analyzed each collection separately using the same

223   settings to calculate effective population size estimates for each collection. For the first analysis

224   that grouped all larval and juvenile collections (N=1815), we also used the adult collections (i.e.,

225   N=1765 for MISC_Adult and Elwha_pbt; Table 1) as candidate parents of unknown sex and

226   assumed that the probability that any of these candidates were true parents was 25%. This prior

227   was chosen as our best estimate of the probability that the actual parent of an offspring in the

228   offspring sample is included in the candidate parents sample as suggested in the manual. We

229   estimated we had obtained a sample of candidate parents that were no more than a quarter of the

230   fish that spawned.

231   SingleSEQUOIA (https://github.com/delomast/singleSequoia, Thomas Delomas, IDFG)

232   is an R package (R Development Core Team 2008) that interfaces with IDFGEN objects to build

233   SEQUOIA (Huisman 2017) inputs for single parent assignment. It also has a function to run

234   simulated analyses in SEQUOIA to estimate expected error rates for single parent assignment

235   given a baseline, marker panel, and number of expected offspring (Supplemental Materials). The

236   threshold for accepting single parent assignments was set to a value of 0.5 log10-likelihood ratio

237   (LLR) between a parent-offspring relationship versus unrelated which was chosen based on prior

238   work that demonstrated this threshold provided an optimal tradeoff between false positive and

239   false negative errors. The results from the simulations confirmed this LLR threshold would yield

240   low error rates for this dataset (Results section).

241   The GSI analyses required two steps to process the reference collections into GSI

242   baselines: 1) constructing reporting groups and 2) testing baseline accuracy. Our goal was to

243   construct an optimal set of reporting groups that maximized both geographic and genetic

244   cohesiveness. Ideally, dividing reference collections into reporting groups that represent the

245 smallest geographic regions possible is most informative for GSI if the baseline achieves high

246 accuracy. We used STRUCTURE v. 2.3.4 (Pritchard et al. 2000) as a heuristic approach to

247 dividing reference collections into the smallest reporting units for accurate GSI discrimination.

248 We used 291 SNPs (having removed 4 species diagnostic markers) and an initial burn-in of 20

249 000 iterations followed by 200 000 iterations of the Markov chain Monte Carlo (MCMC) method

250 to generate posterior probabilities allowing for admixture for each of 40 trials setting K from 1 to

251 10. To obtain a set of robust individual assignments, a membership coefficient (Q) was

252 calculated for each of K_n clusters, and the Greedy algorithm in CLUMPP (Jakobsson and

253 Rosenberg 2007) was used to find the optimal alignment of clusters across multiple runs using

254 method 2 with 100 repeats. Only the top ten replicate STRUCTURE trials with highest mean

255 ln[Pr(X/K)] were averaged using the Greedy algorithm in CLUMPP. The best alignment

256 configuration was computed by 100 configurations that were examined for the highest pairwise

257 similarity (H') to achieve optimal alignment.

258   We tested GSI baseline accuracy using leave-1-out assignment probabilities estimated

259 with the program GENECLASS2 (Piry et al. 2004), which implemented the Bayesian method of

260 Rannala and Mountain (1997). Following assignment of each individual fish, the estimated

261 assignment probabilities to each reporting group were averaged for each collection of fish. The

262 following two baselines were tested: 1) Range-wide and 2) Olympic Peninsula. The Range-wide

263 Baseline contained 21 reference collections (Hess et al. 2013) divided into five different groups

264 by geographic region. Reporting groups were constructed with guidance from the results of

265 STRUCTURE, while trying to maintain geographic cohesion. The Olympic Peninsula Baseline

266 only contained collections within the Elwha River basin (i.e., ELWHA_EF and ELWHA_ST

267 grouped by stream) and its neighboring tributaries (i.e., OLY_EF grouped by river basin) on the

268 Olympic Peninsula. For leave-1-out tests, a level of 90% correct assignment has been used in

269 fisheries management to indicate that baseline populations have been adequately delineated

270 enough to assign individuals from fishery mixtures (e.g. Smith et al. 2005; Beacham et al. 2006;

271 and Seeb et al. 2007). Neither of the two baselines approached this level of accuracy in correct

272 self-assignments across reporting groups, but the Range-wide baseline performed best.

273   After establishing the Range-wide GSI baseline, we applied it to our objectives to

274 determine the natal-origin of fish based on their assignment probabilities to the reporting groups

275 of the GSI baseline. We performed individual assignments of the following unknown mixtures

276   of fish: 1) ELWHA_EF, 2) ELWHA_ST, 3) OLY_EF, 4) Ocean, and 5) MISC_Adult.

277   Individual assignments to the Range-wide baseline reporting groups were estimated with the

278   Bayesian method implemented in gsi_sim (Anderson et al., 2008), which generated a probability

279   of origin using well-established methods for genetic stock identification (Smouse et al. 1990;

280   Paetkau et al. 1995; Rannala and Mountain 1997). The proportion of fish assigned to the

281   reporting group with highest probability ("best estimate" of reporting group of origin) was

282   calculated for each of the unknown mixtures.

283

284   *Objective 1: Determine the origins of larval and juvenile Pacific lamprey of the Elwha River*

285   *basin and attribute production to translocation versus natural recolonization*

286   We used PA to determine whether larvae and juveniles collected in the Elwha River were

287   progeny of any of the sampled adults (translocated versus volitional migrants). Candidate

288   offspring (mixtures 1 - 4 described for GSI above) were assigned to a parent baseline using

289   parent pairs via the program SNPPIT and single parents via the programs COLONY and

290   SEQUOIA. The parent baseline included volitional migrants in the Elwha River and Olympic

291   Peninsula, "ELWHA_pbt(vol)" N=9, translocated adults used for telemetry in the Elwha River

292   basin, "ELWHA_pbt(trans)" N=42, and a set of negative control adults "MISC_Adult" from the

293   Willamette River (N=1714).

294

295   *Objective 2: Quantify the increase in numbers of effective spawners in the Elwha River and*

296   *determine whether current numbers of spawners are equivalent to undammed basins*

297   We used the effective population size (N_e) estimate from COLONY to determine whether

298   effective spawner size (N_b) has increased with time using two screw trap collections in 2018

299   from Indian Creek (N = 447, stream within the Elwha River basin) and the mouth of the Elwha

300   River (N = 69). These two sites were ideal for this analysis because they had relatively large

301   sample sizes spanning a diversity of sizes of fish that represented multiple brood years. These

302   sites were also ideal because Indian Creek was a site in which larvae had been previously

303   quantified to have increased by 14X over recent years and we could compare whether the trend

304   in N_b was similar to this rate of increase. Further, the mouth of the Elwha River was perfect

305   because it represented the productivity of the entire basin, since all fish had to pass this point

306   before migrating to the ocean no matter their natal stream. N_b is the effective number of

307 spawners during a single breeding event and is approximately equal to the metric N_e when N_b is

308 multiplied by the generation time (Waples, 1989). When we separated out each collection of

309 larvae and juveniles by brood year, we could estimate N_b; however, when we analyzed

310 collections together as putative mixtures of brood years, our estimate of effective spawners was

311 likely a hybrid between N_b and N_e. For consistency we distinguished these latter estimates as N_e.

312 Using our size-proxies for aging, these collections were estimated to represent at least six brood

313 years (2013 – 2018). The collection from the mouth of the Elwha River was split into the

314 following three pairs of consecutive brood years to obtain samples sizes >15 fish per sample:

315 BY2013-2014, BY2015-2016, and BY2017-2018. The collection from Indian Creek was split

316 into four brood years: 2013 – 2016. N_b was estimated for each brood year sample within both

317 sites to examine temporal trends.

318 Finally, we estimated N_e using collections of larvae and juveniles from undammed

319 tributaries to the Strait of Juan de Fuca, ocean-phase lamprey, and adult returns at Willamette

320 Falls. These estimates provided context for our estimates of N_e among collections within the

321 Elwha River. For these comparisons, only collections with sample sizes >30 were used to

322 provide the most robust estimates. To avoid complexities of differing growth rates among

323 tributaries and older life stages, we did not attempt to split each collection by estimated brood

324 year; however, length histograms were generated for each collection to aid interpretation of the

325 results.

326 Our N_b analyses required information on brood year composition of larvae to make

327 accurate comparisons across collections. We used length-based age estimates to determine

328 brood years. Larvae in the study ranged 31 – 152 mm while juveniles ranged 86 – 170 mm,

329 which generally translated to more recent versus older brood years, respectively. However, we

330 also captured larvae and juveniles during a range of dates throughout the year, so we used the

331 following methodology to standardize brood year estimates. We delineated age class intervals

332 based on size by first defining a 1-year age class informed by two data points of known ages (i.e.,

333 from parentage assignments). The length range of 30 – 59 mm was used as a proxy for yearlings,

334 and a linear growth rate of 30 mm/year from year 1 to year 4 was used to delineate the

335 subsequent age classes: 2 (60 – 89 mm), 3 (90 – 119 mm), and 4 (120 – 150 mm). Juveniles of

336 any length were assumed to be age 5 (the median Pacific lamprey age at metamorphosis, Dawson

337 et al. 2015). Collection dates were used in conjunction with observed body length to fine-tune

338   age estimates and a single birthdate of May 1st was assumed for consistency when estimating

339   total age. For example, a larva measuring 30 mm collected on May 5th, 2018 was assumed to

340   have a birthdate on May 1st, 2017 and estimated age of 1 year and 4 days. In cases where a larva

341   was captured between October – December of a given year and had an observed length in the

342   lower half of the size range, we assumed that it had experienced the full growing period of that

343   year and adjusted its birthdate accordingly. For example, a larva measuring 30 mm that was

344   collected on October 1st, 2018 was assumed to have a birthdate on May 1st, 2018 (aged 5 months).

345   In contrast, a larva measuring 59 mm collected on October 1st, 2018 was assumed to have a

346   birthdate on May 1st, 2017 (aged 1 year 5 months). These length-based age-class estimates were

347   verified by instances in which we had full-sibship families that were recaptured across years (we

348   examined whether recaptures from a subsequent year fit into the expected range of body lengths

349   for the subsequent age class).

350

351   *Objective 3: Determine the relative productivity of streams within the Elwha River and its*

352   *regional contribution*

353   We used SA to better understand which streams drive productivity in the Elwha and how

354   this productivity may influence the surrounding region. COLONY results were used to determine

355   the composition of stream-of-origin for larvae and juveniles exiting the mouth of the Elwha

356   River, in the ocean, and at Willamette Falls. The collections of larvae and juveniles from sites in

357   the Elwha River basin were first ordered from upstream to downstream and by collection year

358   within each site. We reconstructed full-sibling families among all individuals in the dataset. We

359   assumed that for each pair of full siblings that the full-sibling family originated nearest to the

360   sibling collected at the most upstream site and was present at least since the earliest collection

361   year of either sibling. For example, a juvenile collected at the screw trap at the mouth of the

362   Elwha River was determined to have a natal origin from Indian Creek if it had a full sibling that

363   was collected in Indian Creek. For lamprey collected from the Elwha River basin, we estimated

364   brood year using length-based age methodology to further characterize the composition of natal

365   origins in each collection. One natal origin site (Indian Creek) was detected in relatively high

366   abundance (N > 15) at the mouth of the Elwha River. To better understand ages at outmigration,

367   we compared the composition of Indian Creek ages (length-based estimates) to those at the river

368   mouth for the same collection year.

369

370

371 **Results**

372 *Power and accuracy of genetic methods*

373   For PA and SA applications, we omitted all but one locus for each LD pair because

374 parentage requires independent loci. There were 138 significant pairs of loci in LD found in four

375 major pooled groups North (A, B) , South (P, Q), Lower Columbia (G2, G3, H, I) and Interior

376 Columbia (J2, J3, K, L; Table S1; Hess et al. 2013) which involved 49 unique loci. One locus per

377 LD pair was dropped for a total of 28 dropped and 21 retained, which left a total of 263

378 independent loci.

379   The 263 SNP loci were chosen to be highly informative for parentage applications across

380 the species range; the average MAF across 16 reference populations (Hess et al. 2013) with N >

381 20 varied 5% - 50% with mean average MAF of 31%. The study dataset (Table 1) was cleaned

382 to remove individuals with low genotyping success (>10% missing loci) and duplicate

383 individuals (different at ≤ 5 out of 263 loci); all loci were genotyped at greater than 70% for this

384 entire dataset with an average genotyping rate of 99%. MAF for this dataset was similar to the

385 Range-wide values (i.e., average MAF across 4 pooled pops ELWHA_EF, ELWHA_ST,

386 OLY_EF, and Elwha_pbt varied 2% - 50% with mean average MAF of 30%). The Range-wide

387 dataset (Hess et al. 2013) was also trimmed in a similar way to reduce failed individuals (n=9,

388 >10% missing loci) and no duplicates were found for 263 SNPs. For the trimmed Range-wide

389 dataset a total of 509 individuals across 21 collections were available for analysis.

390   The 263 SNPs demonstrated high power for single parent assignments based on the

391 simulations with the singleSEQUOIA function: an average of analysis-wide 0.0% Type I errors

392 and 0.1% Type II errors across 40 simulations (average accuracy of 99.9%) for the reference

393 collections (N=509 individuals, 21 populations). These values were equivalent to those

394 generated using the two candidate parent pools (i.e., ELWHA_pbt and MISC_Adult).

395   The two GSI baselines (Range-wide and Olympic Peninsula) were found to have

396 accuracies that were lower than our 90% target, based on self-assignment rates in leave-1-out

397 tests. However, the Range-wide GSI baseline performed best and had greatest potential for

398 utility (Table 2). Weighted average self-assignment rates of 21 collections to five reporting

399 groups of the Range-wide GSI baseline ranged from 54% - 87% (weighted mean of 72%, Table

400 2). STRUCTURE analyses helped guide the construction of reporting groups for the Range-

401 wide baseline, as the mean estimates of ln[Pr(X/K)] increased steeply between values for K of 1

402 – 4 and then plateaued between values of K 5 – 7. These results indicated that structuring among

403 the reference collections showed support for as many as 7 different groups (Figure S1), and the

404 delineation into five geographic regions was compatible with the pattern of STRUCTURE

405 coefficients of ancestry (Table S2). In contrast with the accuracies of the Range-wide GSI

406 baseline, the Olympic Peninsula self-assignment rates to 14 groups were lower (Table S3) and

407 ranged from 0% - 81% (weighted mean 31%). STRUCTURE analysis results were not helpful

408 for guiding delineation of groups in the Olympic Peninsula baseline as the estimates of

409 ln[Pr(X/K)] continued to increase with increasing K values of 1 – 6 without any apparent

410 correspondence between coefficients of ancestry and geographic regions. Pooling the Olympic

411 Peninsula baseline into two groups, increased self-assignment success to 64% and 73% for

412 collections inside and outside the Elwha River, respectively (weighted mean 67%, Table S4).

413

414 *Objective 1: Natal origins of Pacific lamprey in the Elwha River*

415 Our first objective was to use PA, SA, and GSI to determine the origins of larvae and juveniles

416 of Pacific lamprey collected in the Elwha River and thereby attribute production to translocation

417 versus natural recolonization. Results using the PA method supported our expectation that

418 production was primarily driven by natural colonization, as evidenced by a lack of parent

419 assignments to translocation parents. We had expected that natural colonization would be a

420 major source of productivity in the Elwha River given the relatively modest translocation effort

421 (<50 adults) to date and because potential sources of Pacific lamprey spawners were observed in

422 nearby basins. The other two methods, SA and GSI, were not adequate for addressing this

423 question. SA could not be used to identify sources of fish from outside the Elwha River and GSI

424 was not precise enough to determine the basin-of-origin for adult colonizers.

425 There were no parent-offspring trios detected in our dataset. This was not unexpected

426 given the small number of candidate parents that were genotyped (N=51. However, there were 2

427 parents that were assigned as single-parents based on COLONY (both adults were volitional

428 migrants collected in the Indian Creek screw trap in 2017) and one of these assignments was

429 concordant with SEQUOIA. Based on our single-parentage power analysis, we would expect

430 only a few Type II errors (0.1% of 1815 candidate offspring; ~2 offspring if all 1815 candidates

431 were true offspring of parents in the baseline) but would not expect false positives. This

432 expectation fits the small discordance with COLONY (i.e. putative single false negative

433 observed). Both parents were collected in Indian Creek in the summer (June and July) of 2017

434 and the two larvae were both collected in October of 2018 from the Elwha River screw trap (site

435 #44), which make their approximate ages around 1 year (1.2 and 1.3 years between collection

436 dates). The larvae measured 40 and 52 mm respectively and would have grown 30 and 38 mm in

437 a year assuming larvae are 4 mm at hatch (Yamazaki et al. 2003).

438 Given the limited parent assignments we could only show direct evidence of reproductive

439 success for volitional adults, but not for translocated adults. The Range-wide GSI baseline

440 showed that collections of larvae and juveniles from the Elwha River (ELWHA_EF and

441 ELWHA_ST) and those from the surrounding Olympic Peninsula (OLY_EF) were most similar

442 to the reporting group representing the regions of Vancouver Island, Puget Sound, and Lower

443 Columbia (VIPSLC, Fig. 3). This level of geographic resolution did not provide informative

444 results for the current objective. The finer geographic scale of the Olympic Peninsula GSI

445 baseline showed some potential for distinguishing the Elwha River basin versus other river

446 basins but with moderate accuracy according to self-assignments (Table S4). The Little River

447 (site 36) was founded by some translocated larvae from Salt Creek and did have one of the

448 highest rates of assignment to "outside" the Elwha River (68%). However, high "outside"

449 assignment rates were also found in Campground Creek (75%, site 16), which was a side channel

450 washed away the year before the collection was made. The Campground Creek site was located

451 relatively far upstream in the upper basin of the Elwha River and would not be expected to have

452 relatively high rates of influence from the outside compared to lower tributaries, but the recent

453 washout may have influenced the genetic composition.

454 The sibship simply determined that spawning likely occurs semelparously and adults that

455 successfully spawn in one basin do not appear to spawn again in adjacent basins. Hence, the

456 objective to determine whether translocation versus volitional or inside versus outside of Elwha

457 River migration occurs was largely inconclusive.

458 We observed full-sibship among fish within the six basins that were represented in our

459 dataset (Elwha River, Lyre River, East Twin River, West Twin River, and Deep Creek), however

460 there were no full siblings shared between basins and no full-sibship among or within the ocean

461 collections (Fig. 2). Within the Elwha River basin, there were a total of 36 full-sibling families

462   (366 individuals represented the 36 families) that were recaptured across space or time or both

463   space and time.

464   We were able to test for positive somatic growth relationships for the 26 full-sibling

465   families that were recaptured across different years (i.e., testing whether individuals from the

466   same family get larger over time). There were 13 positive and 13 negative growth relationships

467   for these recaptured families, and therefore did not appear to suggest consistent growth occurs

468   between full-sibling recaptures. However, when we examined the five sibship families that

469   contained at least 3 or more "recaptures" from each of the collection years, we found that four of

470   the five families showed positive growth (Fig. S2a). Further, growth appeared to decrease with

471   increasing age when we plotted estimated age at first capture versus the estimated growth rate

472   (Fig. S2b). This relationship may explain why full-sibling families recaptured across different

473   years do not always show positive growth. Further, it may indicate that length-based age

474   estimates will be less accurate for older ages due to increasing violation of our assumption that

475   growth rate is constant throughout larval life.

476

477   _Objective 2: Quantify effective spawners in the Elwha River_

478   Our second objective was to quantify the numbers of Pacific lamprey spawners in the

479   Elwha River and determine whether current numbers of spawners have reached capacities

480   observed in neighboring undammed basins. Our expectation was for an increase of effective

481   spawners (N_b) that occurred as rapidly as the increase in larvae observed in a tributary (Indian

482   Creek) of the Elwha River in the five years following dam removal. The larvae captured in the

483   screw trap at Indian Creek increased 14 fold (from 132 larvae in 2014 to 1,805 larvae in 2016;

484   Moser & Paradis 2017). Similarly, we estimated that N_b increased 12 fold across the span of at

485   least four brood years (2013 – 2016) that were represented in our collections of larvae and

486   juveniles from the Indian Creek screw trap in 2018. Further, N_b measured at the mouth of the

487   Elwha River (representative of the overall abundance of the basin) also increased over a five year

488   period to similar values of N_b estimated across neighboring undammed tributaries of the Olympic

489   Peninsula which indicated robust recolonization has likely occurred in the Elwha River basin.

490   The N_e sizes were generally smallest within the Elwha River (average 92, range 5 – 256),

491   larger in other basins of the Olympic Peninsula (average 189, range 124 – 264), and larger still in

492   the ocean-phase collections (average 644, range 461 – 755) and Willamette Falls adult

493   collections (average 1374, range 843 – 1875; Fig. 4). The one exception was the electro-fished

494   and translocated collection from Little River (site #36), which was 256 N_e (95% CI: 149 – 805).

495   The trends in N_b for the screw trap collections grouped by brood year were positive for

496   both the mouth of the Elwha River and Indian Creek (Fig. 5). The oldest brood year collection

497   (spawn year "SY" 2013-2014) at the mouth of the Elwha River had 14 N_b (95% CI: 7 – 32) and

498   increased by 11X to 160 N_b (95% CI: 70 – ∞) in a five year period (Fig. 5). The oldest brood

499   year (SY2013) in Indian Creek was 5 N_b (95% CI: 2 – 20) and increased by ~12x by SY2015 and

500   stabilized at 60 N_b (95% CI: 36 – 122) through SY2016 (Fig. 5). The values observed for the

501   most recent brood years were not significantly smaller (95% confidence values overlapped) than

502   the levels in the undammed basins of the Olympic Peninsula. Moreover, the mean estimate of

503   the most recent brood years of the Elwha River (BY2017-2018, 160 N_b) was in range with the N_e

504   estimates of the undammed basins (Deep Creek, East Twin and West Twin rivers ranged 124 –

505   264 N_e). The N_e estimates of the undammed basins were more akin to N_b estimates because the

506   Olympic Peninsula collections were primarily comprised of large juveniles representing a single

507   brood year (BY2012, Fig. S3). Therefore, our results demonstrated that the most recent brood

508   year N_b estimate for the Elwha River has increased over a five year period to levels similar to

509   that of N_b for undammed tributaries.

510

511   *Objective 3: Relative productivity of Elwha River Pacific lamprey*

512   Our third objective was to determine the relative productivity of streams within the

513   Elwha River and how overall productivity originating from this system is distributed across the

514   broader surrounding region. We expected that within the Elwha River, Indian Creek would be a

515   relatively large source of production given its demonstrated high productivity for salmonids in

516   this basin. Indian Creek, the first tributary upstream from the former Elwha Dam, is an

517   extremely productive stream based on rapidly increasing juvenile production of several

518   salmonids since dam removal (i.e., Chinook Salmon *Oncorhynchus tshawytscha*, Coho Salmon

519   *O. kisutch*, and Steelhead *O. mykiss*; McHenry et al. 2020). Our study, using SA methods, found

520   that this stream was also a primary source of larval and juvenile production of Pacific lamprey

521   for the Elwha River basin, as evidenced by the high (i.e., ~41%) percentage of Indian Creek fish

522   detected in the Elwha River mouth screw trap.

523   Finally, we predicted the contribution of Pacific lamprey from the Elwha River to the

524   broader surrounding region would be relatively small given its recent founding and small area

525   relative to surrounding sources of Pacific lamprey. No emigration from the Elwha River could

526   be detected in neighboring streams or in coastal ocean-phase collections using PA and SA

527   methods. While this may indicate the relatively small contribution the Elwha River makes to

528   panmictic coastal stocks, it may also be an indication of migratory patterns of Pacific lamprey

529   (Clemens et al. 2019). GSI demonstrated that genetic composition of ocean-phase lamprey was

530   reflective of latitudinal stock structure. Each of the genetic methods we used provided

531   complementary information that is key to management and our approach could be useful for

532   other highly dispersive species.

533   We estimated distance traveled for the 32 full-sibling families that were recaptured across

534   space (individual N=337). Among the electrofishing sites, the largest movements were > 12 km

535   downstream from Brannon, Campground, and Hughes creeks (sites #15, #16, #34) to sites near

536   the river mouth (sites #30 and #44). The sites with the greatest geographic diversity of natal

537   origins for full-sibling families were near the river mouth where fish exit to the Strait of Juan de

538   Fuca. Thus, using the sibship results, we were able to trace large portions of the Elwha River

539   productivity based on composition of natal origins of the collections obtained at the river mouth

540   screw trap. In fact, nearly half (N=37, 49%) of all fish in the 2018 collection at the Elwha River

541   mouth screw trap were identified from full-sibling families that were previously sampled

542   upstream from the screw trap (6 unique sites), and most of those fish originated from Indian

543   Creek (N=31, 41%; Fig. 6). The remaining 51% of fish in this 2018 Elwha River mouth screw

544   trap collection were unassigned because not all full sibship families represented in this sample

545   had been previously encountered in the sampling efforts.

546   Interestingly, the composition of ages (estimated based on length and life stage) at the

547   Indian Creek screw trap in 2018 were weighted toward young ages (i.e., brood years 2013, 2014,

548   and 2015 were 21%, 27%, and 46%, respectively); however, among those identified as Indian

549   Creek origin at the river mouth in 2018, they were weighted oppositely (i.e., 80%, 16%, and 4%

550   for brood year 2013, 2014, and 2015, respectively; Fig. 6). This reverse weighting in

551   composition of ages may indicate that either fish tend to reside in the main stem Elwha River

552   before emigrating from the river, or that there is dramatically different size selection between

553   screw traps (i.e., the Elwha River mouth screw trap selects for larger fish). The evidence points

554 to the former explanation because the overall age composition in 2018 across all stream-of-

555 origins at the Elwha River mouth trap was relatively even across brood years similar to the

556 Indian Creek trap (i.e., Elwha River mouth brood year composition was 29%, 9%, 25%, 14%,

557 17%, and 6% for brood years 2013-2018, respectively). The wide length frequency distribution

558 in the Elwha River screw trap in 2018 (Fig. S3) indicates a high degree of age overlap among

559 fish that exit the river, even though there may be a period of residence in the Elwha River main

560 stem.

561   Sibship was not shared between any collections outside of any basin; therefore, we could

562 not use sibship to make any determinations of Elwha River-origin fish outside of the Elwha

563 River. Therefore, we relied on GSI applications for broader regional observations. On average,

564 the individuals in the Elwha dataset across the four pools assigned at a probability of 70% to the

565 most proximate reporting group (i.e., VIPSLC, probability range of 63 – 82%; Fig. 3). The adult

566 collection from Willamette Falls was expected to assign with greatest probability to the

567 Willamette River / Bonneville Pool (WILBON) reporting group and, in fact, 51% assigned to

568 WILBON (Fig. 3). The remaining assignment probabilities for the Willamette Falls adult

569 collection were shared among the following three reporting groups ordered least to greatest

570 probability: Southern U.S. West Coast, SOUTUS; VIPSLC; and interior Columbia River,

571 INTCOL (assignment probabilities of 12%, 16%, and 21%, respectively). The relatively high

572 assignment probabilities across multiple reporting groups is a characteristic that was shared with

573 the Willamette Falls reference collections (Table 2), which showed leave-1-out assignment

574 probabilities from least to greatest to SOUTUS, VIPSLC, INTCOL, and WILBON (8%, 12%,

575 23%, and 54%, respectively).

576   The assignment probabilities of the unknown ocean-phase mixture showed correlation

577 with geography. For example, the assignment probabilities were inversely proportional with

578 latitude of the ocean-phase collections for the SOUTHUS and proportional with latitude for the

579 INTCOL reporting group. The reporting group that includes Elwha River and other Olympic

580 Peninsula origins, VIPSLC, showed the highest assignment probability (40%, Fig. 3) in the

581 ocean-phase collection (site 54, Fig. 1) most proximal to these northern Washington sites. A

582 range of modest assignment probabilities to VIPSLC of 32 – 37% were observed in the ocean-

583 phase collections (sites 50 – 53, Fig. 1) distributed further south to latitude 43.0 N (Fig. 3). The

584 assignment probabilities to VIPSLC decreased to 22% and 0% (Fig. 3) for the ocean-phase

585   collections at the most extreme southern sites 49 and 48 (Fig. 1), respectively. These results

586   suggest there could be relatively broad ocean distribution of the VIPSLC (surrogate for Olympic

587   Peninsula origins) from latitudes 48.0 N to 43.0 N, with decreasing abundance in the extreme

588   southern end of the ocean-phase survey. This broad distribution contrasts with the more

589   restricted INTCOL reporting group, which was found with average assignment probability of

590   >20% (Fig. 3) in just the two northernmost ocean-phase collections (sites 53 and 54, Fig. 1).

591

592

593   **Discussion**

594   Using genetic methods, we were able to demonstrate that adult Pacific lamprey in the

595   Elwha River were successfully reproducing based on direct evidence from parentage, were

596   increasing in abundance through time at rates similar to the observed rate of increase in larval

597   abundance, were producing exceptionally high numbers of offspring from one source stream

598   known for its high productivity of salmonids, and were likely contributing to heterogeneous

599   mixtures of Pacific lamprey in the coastal ocean.

600   Dam removals provide the opportunity for natural recolonization of Pacific lamprey and

601   other anadromous fishes to previously inaccessible upstream habitat. In the Elwha River

602   drainage, our applications of PA, SA, and GSI provided multiple insights into the process of

603   natural recolonization and into the species' biology in general.

604

605   *Production by volitional migrants versus that of translocated adults*

606   When habitat becomes accessible after dam removal or habitat improvements, it is useful

607   to assess the relative merits of taking a passive approach to lamprey recolonization (relying only

608   on volitional migrants) versus actively translocating lamprey to repopulate the newly available

609   habitat (Reid and Goodman 2020). Due to the modest numbers of adults translocated in our

610   study and the limited years available to monitor any translocation production, we could not fully

611   assess the relative merits of these alternative approaches. However, we did document direct

612   evidence of successful spawning of two volitional adults. This suggests that with greater

613   sampling it would be possible to track reproductive efforts of both volitional and translocated

614   fish on a larger scale. Compared to volitional migrants, it is sometimes more feasible to sample

615   candidate parents that are translocated into a system (Hess et al. 2015). For example, in the

616   Umatilla River basin, hundreds of adult Pacific lamprey are genetic sampled and released onto

617   spawning habitats every year (Close et al. 2009). However, our smaller study permitted release

618   of < 50 translocated adults and none of these fish were detected as parents among the collections

619   of larvae that were obtained. While it is likely that this was due to the small sample, it is also

620   possible that these fish were not detected for other reasons. One possibility is that most

621   translocated adults were released in 2018, producing younger, smaller larvae that were difficult

622   to collect. All the translocated lamprey were also radio-tagged, which may have altered their

623   spawning success. If more adults are translocated and sampled, this method could be effective

624   for determination of the relative contributions of translocated fish to existing populations.

625

626   *Increasing trends in N_b signify a robust recolonization and general use of N_b as a proxy for*

627   *abundance*

628   We found that effective spawner size (N_b) within collections in the Elwha River basin

629   was positively correlated with brood year. This was consistent with our expectation for an

630   increase in N_b over time, given the increased abundances of adults, larvae, and juveniles that

631   were observed immediately after the Elwha Dam was removed (Moser & Paradis 2017).

632   Therefore, results from this study supported the general use of N_b as a proxy for characterizing

633   trends in relative spawner abundance. Further, the use of N_e to estimate effective population size

634   also appeared to reflect expected differences in abundances among the samples we compared.

635   That is, the collections from streams within the Elwha River basin showed relatively low N_e

636   compared to undammed tributaries surveyed in the Olympic Peninsula. This result is consistent

637   with the expectation that streams within a recently recolonized tributary will be low in N_e relative

638   to more established populations. All the tributary collections showed lower N_e than ocean phase

639   collections, because ocean phase lamprey would likely be mixtures of natal origins from multiple

640   tributaries, inflating N_e.

641   Field observations helped to validate N_b as a proxy for relative spawner abundance. For

642   example, in Madison Creek the first two lamprey spawners were observed in 2016. Our sibship

643   reconstruction of electro-fished larvae (estimated from brood year 2016) in Madison Creek

644   identified just three full sibship families, 5 unique parents, and estimated an N_b of 5 (95% CI: 2 –

645   20) which closely matched the low number of spawners the field biologists were able to observe.

646   A second example was found in the Indian Creek rotary screw trap where there were no lamprey

647   observed until 2013 when nine adults were captured, and the first larvae were observed in 2014

648   (Moser & Paradis 2017). The genetic analysis of the juveniles captured (estimated from brood

649   year 2013) and sampled in the Indian Creek screw trap in 2018 identified 9 full sibship families,

650   16 unique parents, and estimated an N_b of 5 (95% CI: 2 – 20). From these examples, it appeared

651   that the field biologists were able to observe at least half of the total adults that were estimated to

652   have successfully spawned in the Madison and Indian Creek sites. Given the difficulty observing

653   this cryptic species, this is not a surprising result. The consistent relationship of the field

654   observations and the genetic estimates (genetic estimates of unique numbers of spawners were

655   approximately twice the number observed in the field) helps to validate these results.

656

657   *Relative productivity in the Elwha River*

658   Our finding that Indian Creek is a primary source of production for Pacific lamprey is

659   important because it confirms that stream characteristics that are suitable for salmonid

660   production deliver similar benefits for Pacific lamprey. This evidence is helpful habitat

661   improvement projects targeting multiple species, but with limited budgets. Our results also

662   provided further evidence of how N_e can act as a useful proxy for abundance. In addition to

663   evidence that Indian Creek was a primary source of abundance for the Elwha River (i.e., high

664   proportion of Indian Creek siblings observed at the Elwha River mouth screw trap), we estimated

665   that Indian Creek had relatively high N_e compared to other tributaries. For example, Madison

666   Creek production was not observed at all at the Elwha River mouth and had low N_e.

667

668   *Distribution of natal origins in the ocean-phase*

669   There are many critical uncertainties concerning the marine phase of Pacific lamprey,

670   which include characterizing the distribution patterns of the various source stocks (Clemens et al.

671   2019). Although we could not use PA and SA to precisely estimate the contribution of the

672   Elwha River productivity to the broader region, GSI estimation may prove useful for

673   characterizing the marine distributions of source stocks at a broad spatial scale. In our study, the

674   ocean phase collections showed a decreasing proportion of the group that Elwha River and

675   Olympic Peninsula was most similar to (VIPSLC = Vancouver Island/Puget Sound/Lower

676   Columbia GSI reporting group) from the northern to southern latitudinal collections (40% to 0%,

677   respectively). Consistent with the low but detectable regional patterns of isolation-by-distance

678 gene-flow of Pacific lamprey reported by Spice et al. (2012), our results suggest evidence of
679 relatively restricted movement compared to a state of panmixia. For example, although the
680 VIPSLC stock is relatively broadly distributed along the coast, it appears to concentrate near
681 natal rivers. In general, Pacific lamprey stocks appeared as heterogenous mixtures in the ocean,
682 reflective of the geographic distribution of source stocks. This suggests that anadromous lamprey
683 will occur most frequently in the ocean regions proximate to their natal river. However, our
684 results indicated that marine patterns may depart from this basic assumption because the
685 distribution of the interior Columbia River (INTCOL) reporting group was weighted north in the
686 ocean survey rather than uniformly distributed around the mouth of the Columbia River. Finally,
687 we found that there were relatively high probabilities of assignment to multiple reporting groups
688 among adult fish in the collections from Willamette Falls. This suggests that Willamette Falls
689 may harbor exceptional diversity of stock origins relative to the other areas that were sampled.
690   At finer scales within the Olympic Peninsula, GSI may be able to discern Elwha River-
691 origin Pacific lamprey from all other undammed tributaries of the Olympic Peninsula. However,
692 this application is of limited utility because there are few places in which the Elwha River and
693 these other Olympic Peninsula stocks occur as a mixed stock (e.g., Strait of Juan de Fuca).
694 Further, the low degree of self-assignment accuracy for finer scale discrimination of different
695 Olympic Peninsula tributaries diminishes the potential utility of GSI at fine spatial scales.
696   Although full-sibling families were detected between years and sites within basins, the
697 occurrence of full-siblings did not extend outside tributaries. This result supported the
698 semelparous biology of the species: after adults return to freshwater, they typically lack the
699 ability to repeat spawn in neighboring tributaries. The high degree of spatial autocorrelation (i.e.,
700 lack of full-sibship detections between basins and presence of full siblings shared across time
701 and space within basins) provides support for the accuracy of the SNP panel for pedigree
702 reconstruction given the expectation that relatedness would be spatially auto-correlated.
703 However, the lack of full-sibship sharing between any of the tributary collections and the ocean
704 phase collections suggests that our sample sizes were insufficient to overcome low detection
705 probabilities, given the scale of mixed natal origins of all the tributaries represented by ocean
706 collections. This low detection probability is one limitation of PA and SA, particularly if the
707 objective is to identify natal origins on a broad spatial scale (e.g., ocean-phase mixed stock).
708

709 *Use of size and life stage as proxies for age class in young populations of Pacific lamprey*

710   Accurate aging of larval and juvenile Pacific lamprey is a difficult challenge for fisheries

711 managers. In bony fishes, otoliths and scales are useful anatomical features to derive ages

712 (Campana 2001). Lampreys lack these structures and statoliths are imperfect substitutes for

713 otoliths (Potts et al. 2015). The use of lengths to derive ages has been used in other studies

714 (Beamish and Medland 1988), albeit with increased severity of limitations for older ages (>5

715 years). While PA can be an ideal method for aging (Hess et al. 2015), there were not enough

716 parents sampled in this study. However, PA was useful for ground truthing several size classes

717 specific to the Elwha River basin, but we realized there were likely increasing violations of our

718 assumed constant growth rates with increasing age. From our validation testing with sibship

719 recaptures, our crude length-age relationship served as a useful proxy for age and to consistently

720 separate brood years. Further, similar length-age relationships could be useful in other study

721 populations if they were founded within a short span of years as in the Elwha River (e.g. less

722 than five). However, systems that have achieved equilibrium for a long period could host

723 enormous overlap in lamprey ages based on the protracted freshwater phase of larval lampreys

724 (e.g., sea lamprey larvae can persist for 0-12+ years, Potter 1980), rendering length-based age

725 estimation worthless.

726

727 *Advancements in technology*

728   Advancements in technology have enabled implementation of genetic methods (PA, SA,

729 and GSI) to determine natal origins for highly dispersive species. The feasibility of genotyping

730 the high number of markers required by these methods is possible due to laboratory

731 advancements. For example, use of genotyping-in-thousands by sequencing (GT-seq, Campbell

732 et al. 2015) helped to overcome cost-efficiency challenges in this study.

733   The selection of loci was also key to our success. Making the right choices required

734 striking a balance among multiple objectives for management; SNPs that are ideal for discerning

735 population level differences (i.e. good GSI candidates) are often relatively poor in

736 informativeness for PA and SA across the species range (Hess et al. 2015). In this study, the loci

737 selected showed utility for GSI at broad regional scales when using all available SNPs together

738 (291, including all high F_{ST} SNPs). We also employed a robust subset of these markers (i.e., 263

739 independent, high minor allele frequency SNPs) that met the high-informativeness requirements

740   for accurate single parentage analysis and sibship reconstruction. Together, these

741   complementary methods allowed us to achieve our objectives and identify both regional and

742   fine-scale patterns of relatedness for Pacific lamprey within the Elwha River basin, across the

743   local area in the Olympic Peninsula, and in the coastal ocean.

744   All of the methods we tested had useful attributes. Management objectives requiring fast

745   turnaround and high precision for natal-origin estimation may benefit from a focus on SA

746   baseline sampling. However, when candidate-parent sampling challenges can be overcome to

747   achieve high rates of parent sampling, PA methods are advantageous because both age and natal-

748   origin can be accurately estimated (Steele et al. 2019). Finally, in cases where sampling

749   resources are more limited and the high level of precision is not needed, GSI is likely the most

750   cost-effective method to determine natal origins.

751

752   *Conclusion*

753   For a highly dispersive fish species like Pacific lamprey, these analyses, methods, and

754   informative molecular markers can help illuminate a myriad of management and biological

755   questions that had been stymied by the low population genetic structure of the species. Our

756   results indicate that with more extensive sampling, it would be possible to estimate larval

757   recruitment from various management units, be they translocated adult lamprey, artificially

758   propagated larvae, or production from individual drainage basins. Understanding sources and

759   sinks of lamprey production has been an elusive goal and is critical to recovery of species at risk

760   (Lucas et al. In Press). Genetic methods and more extensive sampling could put this goal within

761   reach and shed light on a plethora of unknowns in lamprey biology (e.g., duration of ocean phase,

762   oceanic distribution and movements, life history variation, etc.).

References

775 **References**

776 Ackerman, M. W., Hand, B. K., Waples, R. K., Luikart, G., Waples, R. S., Steele, C. A., Garner,

777 B.A., McCane, J., and Campbell, M. R. (2017). Effective number of breeders from sibship

778 reconstruction: empirical evaluations using hatchery steelhead. Evolutionary applications, 10(2),

779 146-160.

780 Anderson, E. C., Waples, R. S., and Kalinowski, S. T. 2008. An improved method for estimating

781 the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences,

782 65: 1475–1486.

783 Anderson, Eric C. "Large-scale parentage inference with SNPs: an efficient algorithm for

784 statistical confidence of parent pair allocations." Statistical applications in genetics and

785 molecular biology 11.5 (2012).

786 Arakawa, H., Lampman, R. T., & Hess, J.E. (in review). Who's kid did you eat? Genetic

787 identification of species and parents of larval lampreys in fish predator guts. North American

788 Journal of Fisheries Management.

789 Araujo, H. A., Candy, J. R., Beacham, T. D., White, B., & Wallace, C. (2014). Advantages and

790 challenges of genetic stock identification in fish stocks with low genetic resolution. Transactions

791 of the American Fisheries Society, 143(2), 479-488.

792 Beamish, F. W. H., & Medland, T. E. (1988). Age determination for lampreys. Transactions of

793 the American Fisheries Society, 117(1), 63-71.

794 Campana, S.E. 2001. Accuracy, precision and quality control in age determination, including a

795 review of the use and abuse of age validation methods. Journal of Fish Biology 59: 197–242.

796 Campbell, M. R., Kozfkay, C. C., Copeland, T., Schrader, W. C., Ackerman, M. W., & Narum, S.

797 R. (2012). Estimating abundance and life history characteristics of threatened wild Snake River

798 steelhead stocks by using genetic stock identification. Transactions of the American Fisheries

799 Society, 141(5), 1310-1327.

800 Campbell, N. R., Harmon, S. A., & Narum, S. R. (2015). Genotyping-in-Thousands by

801 sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon

802 sequencing. Molecular ecology resources, 15(4), 855-867.

803 Candy, J. R., Campbell, N. R., Grinnell, M. H., Beacham, T. D., Larson, W. A., & Narum, S. R.

804 (2015). Population differentiation determined from putative neutral and divergent adaptive

805   genetic markers in Eulachon (Thaleichthys pacificus, Osmeridae), an anadromous Pacific smelt.

806   Molecular Ecology Resources, 15(6), 1421-1434.

807   Clabough, T.S., Keefer, M.L., Caudill, C.C., Johnson, E.L. and Peery, C.A., 2012. Use of night

808   video to enumerate adult Pacific lamprey passage at hydroelectric dams: challenges and

809   opportunities to improve escapement estimates. North American Journal of Fisheries

810   Management, 32(4), pp.687-695.

811   Clemens, B. J. (2019). A call for standard terminology for lamprey life stages. Fisheries, 44(5),

812   243-245.

813   Clemens, B. J., Weitkamp, L., Siwicke, K., Wade, J., Harris, J., Hess, J., ... & Orlov, A. M.

814   (2019). Marine biology of the pacific lamprey Entosphenus tridentatus. Reviews in Fish Biology

815   and Fisheries, 1-22.

816   Close DA, Currens KP, Jackson A, Wildbill AJ, Hansen J, Bronson P, Aronsuu K (2009)

817   Lessons from the reintroduction of a noncharismatic, migratory fish: Pacific lamprey in the upper

818   Umatilla River, Oregon. In: Brown LR, Chase SD, Mesa MG, Beamish RJ, Moyle PB (eds)

819   Biology, management and conservation of lampreys in North America. Am Fish Soc Symp 72,

820   Bethesda, MD, pp 233–253

821   Dawson, H.A., Quintella, B.R., Almeida, P.R., Treble, A.J. and Jolley, J.C., 2015. The ecology

822   of larval and metamorphosing lampreys. In Lampreys: biology, conservation and control (pp. 75-

823   137). Springer, Dordrecht.

824   Hess, J. E., Campbell, N. R., Close, D. A., Docker, M. F., & Narum, S. R. (2013). Population

825   genomics of Pacific lamprey: adaptive variation in a highly dispersive species. Molecular

826   Ecology, 22(11), 2898-2916.

827   Hess, J.E., Campbell, N.R., Docker, M.F., Baker, C., Jackson, A., Lampman, R., McIlraith, B.,

828   Moser, M.L., Statler, D.P., Young, W.P. and Wildbill, A.J., 2015. Use of genotyping by

829   sequencing data to develop a high-throughput and multifunctional SNP panel for conservation

830   applications in Pacific lamprey. Molecular Ecology Resources, 15(1), pp.187-202.

831   Hess, J. E., Smith, J. J., Timoshevskaya, N., Baker, C., Caudill, C. C., Graves, D., ... & Silver, G.

832   (2020). Genomic islands of divergence infer a phenotypic landscape in Pacific lamprey.

833   Molecular Ecology. https://doi.org/10.1111/mec.15605

834   Hogg R., Coughlan S.M. Jr., and Zydlewski J. 2013. Anadromous sea lampreys recolonize a

835   Maine coastal river tributary after dam removal. Trans. Am. Fish. Soc. 142:1381 1394.

836   Huisman, J. (2017). Pedigree reconstruction from SNP data: parentage assignment, sibship

837   clustering and beyond. *Molecular ecology resources*, *17*(5), 1009-1024.

838   Johnson, N.S., Swink, W.D., Dawson, H.A. and Jones, M.L., 2016. Effects of coded-wire-

839   tagging on stream-dwelling Sea Lamprey larvae. North American Journal of Fisheries

840   Management, 36(5), pp.1059-1067.

841   Jolley, J. C., Silver, G. S., Harris, J. E., & Whitesel, T. A. (2018). Pacific lamprey recolonization

842   of a Pacific Northwest river following dam removal. River research and applications, 34(1), 44-

843   51.

844   Jones, O. R., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from

845   multilocus genotype data. Molecular ecology resources, 10(3), 551-555.

846   Kan, T. (1975). Systematics, variation, distribution, and biology of lampreys of the genus

847   *Lampetra* in Oregon.: Oregon State University.

848   Lucas, M.C., Hume, J.B., Almeida, P.R., Aronsuu, K., Habit, E., Silva, S., Wang, C.,Zampatti, B.

849   (In Press). Emerging conservation initiatives for lampreys:

850   research challenges and opportunities. Journal of Great Lakes Research.

851   McHenry, M., Elofson, M., Liermann, M., Bennett, T., Corbett, S., and Pess, G. 2020. 2019

852   Elwha River smolt enumeration project report. Submitted to Olympic National Park by the

853   Elwha Klallam Tribe February 28, 2020.

854   Moser, M.L., Butzerin, J.M. and Dey, D.B., 2007. Capture and collection of lampreys: the state

855   of the science. Reviews in Fish Biology and Fisheries, 17(1), pp.45-56.

856   Moser ML, Almeida PR, Kemp P, Sorenson PW. 2015. Spawning migration. Chapter 5, pages

857   215–262 in Docker MF, editor. Lampreys: biology, conservation and control. Fish & Fisheries

858   Series 37, Corvallis, Oregon, USA, DOI https://doi.org/10.1007/978-94-017-9306-3_5.

859   Moser, M. L., & Paradis, R. L. (2017). Pacific lamprey restoration in the Elwha River drainage

860   following dam removals. Am Curr, 42, 3-8.

861   Moser, M.L., Jackson, A.D., Mueller, R.P., Maine, A.N. and Davisson, M., 2017. Effects of

862   passive integrated transponder (PIT) implantation on Pacific lamprey ammocoetes. Animal

863   Biotelemetry, 5(1), p.1.

864   Moser, M.L., P. R. Almeida, J. J King, and E. Pereira. In Press. Passage and freshwater habitat

865   requirements of anadromous lampreys: considerations for conservation and control. J. Great

866   Lakes Research. https://doi.org/10.1016/j.jglr.2020.07.011.

867   Nielsen, J. L., Byrne, A., Graziano, S. L., & Kozfkay, C. C. (2009). Steelhead genetic diversity

868   at multiple spatial scales in a managed basin: Snake River, Idaho. North American Journal of

869   Fisheries Management, 29(3), 680-701.

870   Paetkau, D., W. Calvert, I. Stirling, and C. Strobeck. 1995. Microsatellite analysis of population

871   structure in Canadian polar bears. Molecular Ecology 4:347–354.

872   Pereira E., Quintella B.R., Mateus C.S., Alexandrade C.M., Belo A.F., Telhado A., Quadrado

873   M.F., and Almeida P.R. 2017. Performance of a vertical‐slot fish pass for the Sea Lamprey

874   Petromyzon marinus L. and habitat recolonization. River Resarch Applications 33:16 26.

875   Potter IC (1980) Ecology of larval and metamorphosing lampreys. Can J Fish Aquat Sci

876   37:1641-1657

877   Potts, D.D., Dawson, H.A. & Jones, M.L. Validation of a relationship between statolith size and

878   age of larval Great Lakes sea lamprey (Petromyzon marinus). Environmental Biology of Fish 98,

879   1859–1869 (2015). https://doi.org/10.1007/s10641-015-0403-7

880   Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes.

881   Proceedings of the National Academy of Sciences of the USA 94:9197–9201.

882   Reid, S. B. Goodman, D.H. 2020. Natural recolonization by Pacific lampreys in a Southern

883   California coastal drainage: implications for their biology and conservation. N. Am. J. Fish.

884   Mgmt. 40:335–341. https://doi.org/10.1002/nafm.10412.

885   Schultz LD, Chasco BE, Whitlock SL, Meeuwig MH, Schreck CB. Growth and annual survival

886   estimates to examine the ecology of larval lamprey and the implications of ageing error in fitting

887   models. J Fish Biol. 2017;90(4):1305‐1320. doi:10.1111/jfb.13230

888   Scribner KT, Jones ML (2002) Genetic assignment of larval parentage as a means of assessing

889   mechanisms underlying adult reproductive success and larval dispersal. Great Lakes Fishery

890   Commission Completion Report, Ann Arbor, MI.

891   Shaklee, J. B., Beacham, T. D., Seeb, L., & White, B. A. (1999). Managing fisheries using

892   genetic data: case studies from four species of Pacific salmon. Fisheries Research, 43(1-3), 45-78.

893   Smith JJ, Baker C, Eichler EE et al (2012) Genetic consequences of programmed genome

894   rearrangement. Curr Biol 22:1524–1529.

895   Smouse, P.E., R.S. Waples, and J.A. Tworek. 1990. A genetic mixture analysis for use with

896   incomplete source population data. Canadian Journal of Fisheries and Aquatic Sciences 47: 620–

897   634.

898 Steele, C. A., Anderson, E. C., Ackerman, M. W., Hess, M. A., Campbell, N. R., Narum, S. R.,

899 & Campbell, M. R. (2013). A validation of parentage-based tagging using hatchery steelhead in

900 the Snake River basin. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(7), 1046-1054.

901 Steele, C. A., Hess, M., Narum, S., & Campbell, M. (2019). Parentage‐based tagging:

902 Reviewing the implementation of a new tool for an old problem. Fisheries, 44(9), 412-422.

903 Wang JL (2004) Sibship reconstruction from genetic data with typing

904 errors. Genetics, 166, 1963–1979.

905 Wang J (2009) A new method for estimating population sizes from a single

906 sample of multilocus genotypes. Molecular Ecology, 18, 2148–2164.

907 Waples, R.S., 1990. Conservation genetics of Pacific salmon. III. Estimating effective population

908 size. Journal of Heredity, 81(4), pp.277-289.

909 Ward, D. L., Clemens, B. J., Clugston, D., Jackson, A. D., Moser, M. L., Peery, C., & Statler, D.

910 P. (2012). Translocating adult Pacific lamprey within the Columbia River Basin: state of the

911 science. Fisheries, 37(8), 351-361.

912 Yamazaki Y, Fukutomi N, Takeda K, Iwata A (2003) Embryonic development of the Pacific

913 lamprey, Entosphenus tridentatus. Zool Sci 20:1095–1098.

914 Table 1. Locations (latitude, longitude), group names, number of samples (N), and analysis "Type" for each Pacific lamprey collection

915 site. Map site numbers correspond to those in Figure 1.

Map site#	Description (Latitude [°N], Longitude [°W])	Pooled group	Collection Year(s)	N	Type
15	Brannon Creek Rkm 0.11 (48.048 -123.590)	ELWHA_EF	2018	11	L/J Offsp.
16	Campground Creek Rkm 0.2 (48.030 -123.589)	ELWHA_EF	2017	9	L/J Offsp.
29	Elwha River Rkm 0.18 (48.146 -123.568)	ELWHA_EF	2018	2	L/J Offsp.
30	Elwha River Rkm 0.38 (48.146 -123.564)	ELWHA_EF	2018	37	L/J Offsp.
31	Elwha River Rkm 0.9 (48.142 -123.561)	ELWHA_EF	2018	19	L/J Offsp.
32	Elwha River Rkm 11.6 (48.070 -123.577)	ELWHA_EF	2018	1	L/J Offsp.
33	Elwha River Rkm 12.7 (48.060 -123.578)	ELWHA_EF	2018	14	L/J Offsp.
34	Hughes Creek Rkm 0.1 (48.025 -123.595)	ELWHA_EF	2017	8	L/J Offsp.
35	Indian Creek Rkm 0.0 (48.066 -123.579)	ELWHA_EF	2017	18	L/J Offsp.
36	Little River Rkm 4.6 (48.060 -123.519)	ELWHA_EF	2018	38	L/J Offsp.
37	Madison Creek Rkm 0.0 (48.046 -123.589)	ELWHA_EF	2017	51	L/J Offsp.
		ELWHA_EF	*SUBTOTAL*	*208*	L/J Offsp.
44	Elwha River Rkm 0.67 (48.145 -123.562)	ELWHA_ST	2016, 2017, 2018	1, 4, 76	L/J Offsp.
45	Indian Creek Rkm 0.63 (48.067 -123.585)	ELWHA_ST	2016, 2017, 2018	10, 11, 455	L/J Offsp.
46	Lyre River Mouth Rkm 0.0 (48.158 -123.828)	ELWHA_ST	2018	9	L/J Offsp.
		ELWHA_ST	*SUBTOTAL*	*566*	L/J Offsp.
17	Deep Creek Rkm 2.0 (48.160 -124.035)	OLY_EF	2015, 2016, 2017, 2018	47, 20, 74, 27	L/J Offsp.
18	East Twin River Rkm 2.2 (48.149 -123.936)	OLY_EF	2016, 2017, 2018	10, 100, 25	L/J Offsp.
19	West Twin River Rkm 1.9 (48.158 -123.969)	OLY_EF	2017, 2018	32, 11	L/J Offsp.
		OLY_EF	*SUBTOTAL*	*346*	L/J Offsp.

48	Eastern Pacific Ocean US West Coast Slope NLat42SLat41	Ocean	2017	3	J Offsp.
49	Eastern Pacific Ocean US West Coast Slope NLat43SLat42	Ocean	2017	85	J Offsp.
50	Eastern Pacific Ocean US West Coast Slope NLat44SLat43	Ocean	2017	216	J Offsp.
51	Eastern Pacific Ocean US West Coast Slope NLat45SLat44	Ocean	2017	161	J Offsp.
52	Eastern Pacific Ocean US West Coast Slope NLat46SLat45	Ocean	2017	67	J Offsp.
53	Eastern Pacific Ocean US West Coast Slope NLat47SLat46	Ocean	2017	130	J Offsp.
54	Eastern Pacific Ocean US West Coast Slope NLat48SLat47	Ocean	2017	30	J Offsp.
-		Ocean	2017, 2018	1, 2	J Offsp.
		Ocean	*SUBTOTAL*	*695*	J Offsp.
47	Willamette Falls Fish Ladder Rkm 41.1 (45.352 - 122.618)	MISC_Adult	2014, 2015, 2016	865, 580, 269	A Parent
		MISC_Adult	*SUBTOTAL*	*1714*	A Parent
1	Clallam River Rkm 9.2 (48.201 -124.271)	ELWHA_pbt(vol)	2018	1	A Parent
5	Elwha River Rkm 0.38 (48.147 -123.564)	ELWHA_pbt(vol)	2017	1	A Parent
6	Elwha River Rkm 0.67 (48.145 -123.562)	ELWHA_pbt(vol)	2017	1	A Parent
8	Indian Creek Rkm 0.63 (48.067 -123.585)	ELWHA_pbt(vol)	2018, 2017	1, 5	A Parent
		ELWHA_pbt(vol)	*SUBTOTAL*	*9*	A Parent

11	Elwha River- Pedestrian bridge Rkm 5.26 (48.114 -123.554)	ELWHA_pbt(trans)	2017	7	A Parent
12	Elwha River- Sonar Rkm 2.47(48.134 -123.554)	ELWHA_pbt(trans)	2017	1	A Parent
13	Elwha River- Tribal hatchery Rkm 2.77 (48.132 -123.551)	ELWHA_pbt(trans)	2017, 2018	1,25	A Parent
14	Elwha River- WDFW hatchery Rkm 4.63 (48.120 -123.554)	ELWHA_pbt(trans)	2017	8	A Parent
		ELWHA_pbt(trans)	*SUBTOTAL*	*42*	A Parent

916   Note: The analysis "Type" indicates whether the collection was comprised of larvae and juveniles (L/J), juveniles only (J), and Adults

917   only (A) as well as whether the collections were treated as candidate offspring (Offsp.) or candidate parents (Parent) for the different

918   analyses in this study.

919 Table 2. Cross validation matrix for leave-one-out procedure on reference collections in the Pacific lamprey genetic baseline using

920 291 SNP loci. Group definitions as follows: Northern B.C. ("NORTBC), Vancouver Island/Puget Sound/Lower Columbia

921 ("VIPSLC"), Willamette River and Bonneville Reservoir ("WILBON"), interior Columbia River ("INTCOL"), and Southern U.S.

922 West Coast ("SOUTUS").

Pop #	Group	N	Collection	NORTBC	VIPSLC	WILBON	INTCOL	SOUTUS	Avg. Self (Weighted)
Pop01	NORTBC	17	Nass	73.7	0.0	9.1	17.1	0.0	63.9
Pop02	NORTBC	9	Skeena	45.3	0.0	0.2	54.5	0.0	
Pop03	VIPSLC	4	Deadman	0.1	24.6	0.2	75.0	0.0	69.2
Pop04	VIPSLC	30	Stamp	0.0	96.7	0.0	0.0	3.3	
Pop05	VIPSLC	4	Hood	0.0	100.0	0.0	0.0	0.0	
Pop06	VIPSLC	30	Toutle	0.0	43.6	25.9	0.0	30.5	
Pop07	WILBON	30	Willamette 1995	1.3	6.0	46.5	40.0	6.2	53.9
Pop08	WILBON	29	Willamette 2011 Early	2.0	19.2	56.2	13.9	8.7	
Pop09	WILBON	30	Willamette 2011 Late	3.0	11.8	60.7	16.6	7.9	
Pop10	WILBON	29	Klickitat	0.0	9.9	56.3	25.7	7.9	
Pop11	WILBON	30	Fifteenmile	0.0	0.0	50.1	43.0	6.9	
Pop12	INTCOL	30	Deschutes 1995	0.0	0.0	1.3	98.7	0.0	86.5
Pop13	INTCOL	30	Deschutes 2011 Early	1.8	0.0	41.9	53.1	3.3	
Pop14	INTCOL	30	Deschutes 2011 Late	0.0	0.0	20.8	75.9	3.2	
Pop15	INTCOL	30	Wenatchee	0.0	0.0	4.3	95.7	0.0	
Pop16	INTCOL	39	Entiat	0.0	0.0	1.0	99.0	0.0	

Pop17	INTCOL	5	Methow	0.0	0.0	0.0	100.0	0.0	
Pop18	INTCOL	29	Snake	0.1	0.0	9.1	90.8	0.0	
Pop19	SOUTUS	9	Coquille	0.0	3.8	11.3	0.0	84.9	77.4
Pop20	SOUTUS	30	Rogue	0.0	12.4	5.7	2.2	79.7	
Pop21	SOUTUS	35	Klamath	0.0	11.4	14.9	0.1	73.6	
		509							

923 Note: Samples sizes (N) and average self-assignments (Avg. Self) are provided for each reference collection. Collections are

924 composed of individuals analyzed by Hess et al. (2013).
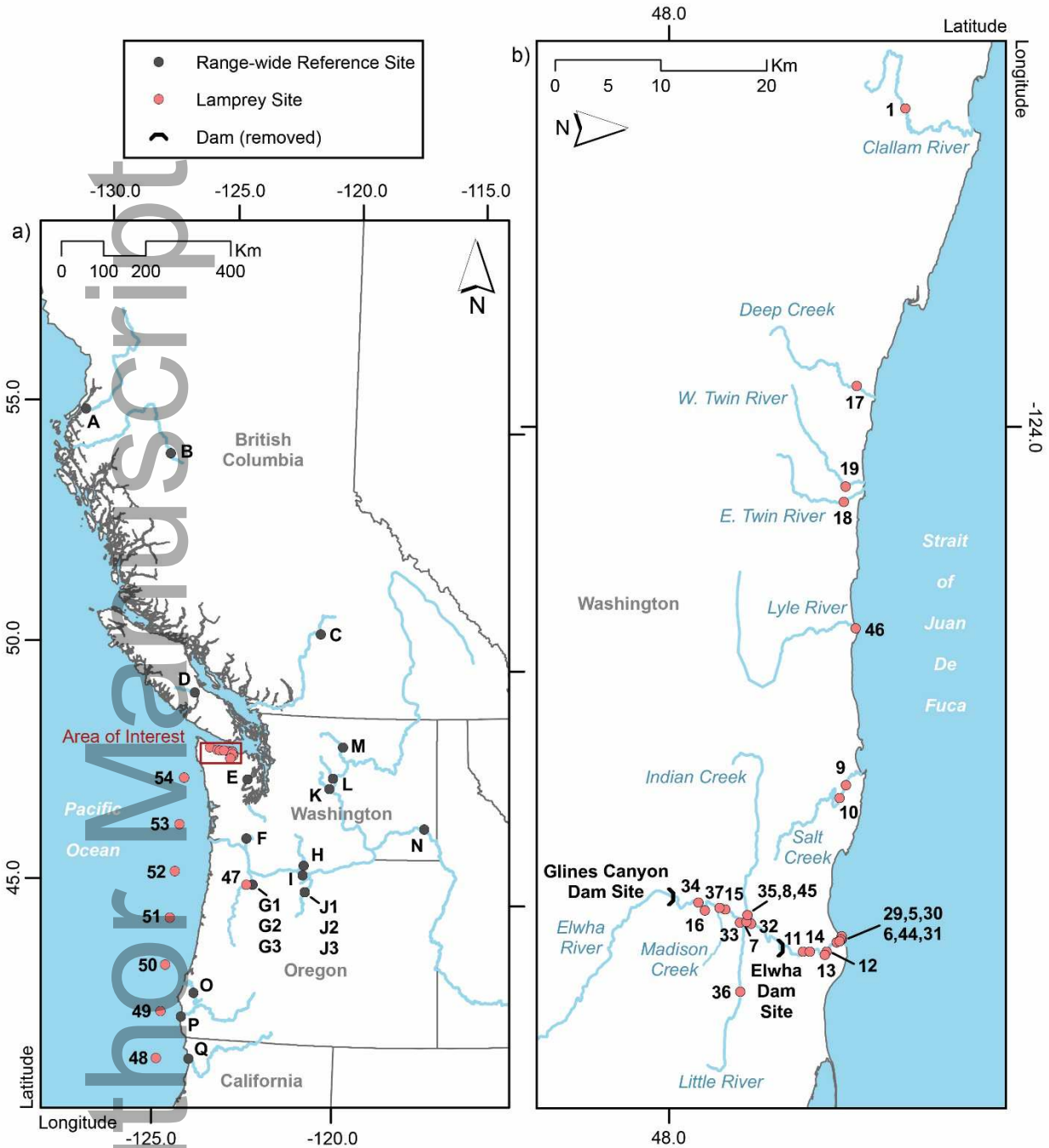
Figure 1. a) Map of the entire study area, box that indicates the area of interest on the Olympic Peninsula, and "Range-wide Reference Sites" (black circles, letter codes as in Table S1). "Lamprey Sites" for Pacific lamprey samples genotyped for this study (red circles) are labeled with the "Map site #s" in Table 1. b) The inset map is the "area of interest" and includes the Elwha River collections and the collections outside the Elwha River distributed across the Olympic Peninsula that were used for context. Two dam sites where dams were removed in the Elwha River are indicated on the inset. Note that the inset map north arrow points right.

Figure 2. Sibship sharing among collections of Pacific lamprey. The matrix below shows sample collections by Year and Site, with column headers indicating collection identifiers.

		2017 EL_hughes_Rkm0001	2017 EL_campgr_Rkm0002	2017 EL_madiso_Rkm0000	2018 EL_branon_Rkm0001	2018 EL_elwhar_Rkm0127	2018 EL_little_Rkm0046	2017 EL_indcrk_Rkm0000	2018 EL_elwhar_Rkm0116	2018 EL_elwhar_Rkm0009	2018 EL_elwhar_Rkm0002	2018 EL_elwhar_Rkm0004	2016 EL_INDSCR_Rkm0006	2017 EL_INDSCR_Rkm0006	2018 EL_INDSCR_Rkm0006	2016 EL_ELWSCR_Rkm0007	2017 EL_ELWSCR_Rkm0007	2018 EL_ELWSCR_Rkm0007	2018 LY_LYRSCR_Rkm0000	2015 DC_deepcr_Rkm0020	2016 DC_deepcr_Rkm0020	2017 DC_deepcr_Rkm0020	2018 DC_deepcr_Rkm0020	2016 ET_etwinr_Rkm0022	2017 ET_etwinr_Rkm0022	2018 ET_etwinr_Rkm0022	2017 WT_wtwinr_Rkm0019	2018 WT_wtwinr_Rkm0019	2017 EP_USWCSL_NLat42SLat41	2017 EP_USWCSL_NLat43SLat42	2017 EP_USWCSL_NLat44SLat43	2017 EP_USWCSL_NLat45SLat44	2017 EP_USWCSL_NLat46SLat45	2017 EP_USWCSL_NLat47SLat46	2017 EP_USWCSL_NLat48SLat47	2017 EP_USWCSL_NLatunkSLatunk	2018 EP_USWCSL_NLatunkSLatunk	N
Year	Site	34	16	37	15	33	36	35	32	31	29	30	45	45	45	44	44	44	46	17	17	17	17	18	18	18	19	19	48	49	50	51	52	53	54	-	-	
2017	34	8																																				8
2017	16	2	7																																			9
2017	37			51																																		51
2018	15		2		9																																	11
2018	33	1		3		10																																14
2018	36						38																															38
2017	35							18																														18
2018	32								1																													1
2018	31		1							18																												19
2018	29										2																											2
2018	30	1		2	1							33																										37
2016	45												6			4																						10
2017	45													4		1	6																					11
2018	45														76		2	40	337																			455
2016	44																			1																		1
2017	44														4																							4
2018	44	2			2			5		1		1	2	1	23			39																				76
2018	46																		9																			9
2015	17																			47																		47
2016	17																			1	19																	20
2017	17																			9	2	63																74
2018	17																			2	1	3	21															27
2016	18																							10														10
2017	18																							9	91													100
2018	18																							1	11	13												25
2017	19																										32											32
2018	19																										1	10										11
2017	48																												3									3
2017	49																													85								85
2017	50																														216							216
2017	51																															161						161
2017	52																																67					67
2017	53																																	130				130
2017	54																																		30			30
2017	-																																			1		1
2018	-																																				2	2
		13	11	54	13	11	38	109	1	19	2	34	9	47	364	1	0	39	9	59	22	66	21	20	102	13	33	10	3	85	216	161	67	130	30	1	2	1815

Elwha_EF Elwha_ST OLY_EF Deep Creek OLY_EF East Twin OLY_EF West Twin Ocean

932 Figure 2. Sibship sharing among collections of Pacific lamprey. Colors indicate the number of full-sibling families that were first

933   captured from their original site (green), first captured in the same year but different location (yellow), first captured in a different year

934   but same site (blue), or first captured in a different year and different site (purple). Site numbers correspond with Figure 1 and Table 1.

935
936 Figure 3. Genetic stock composition of Pacific lamprey samples in this study. Samples include

937 ELWHA_EF, ELWHA_ST, OLY_EF, ELWHA_PBT, MISC_Adult, and Ocean as described in

938 Table 1. The colors correspond to proportions of individuals that assigned to the five reporting

939 groups of the Range-wide GSI baseline: Northern B.C. ("NORTBC), Vancouver Island/Puget

940 Sound/Lower Columbia ("VIPSLC"), Willamette River and Bonneville Reservoir ("WILBON"),

941 interior Columbia River ("INTCOL"), and Southern U.S. West Coast ("SOUTUS") (Table 2)..
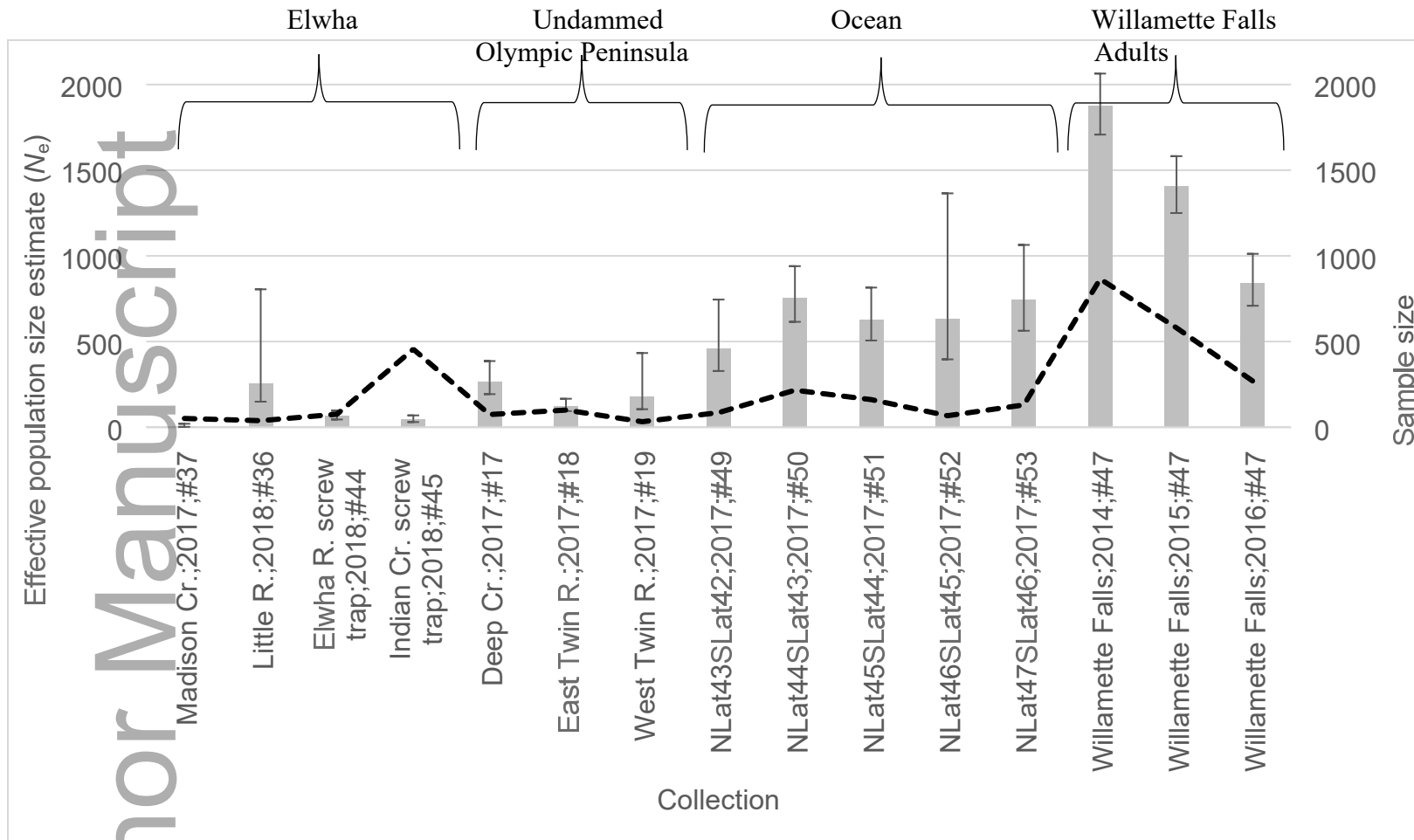
942   Figure 4. Effective population size estimates (N_e) with 95% confidence intervals and sample sizes (dashed line on secondary y-axis)

943   for each Pacific lamprey collection. Collections are grouped into categories according to the region they are from and indicate site

944   names, collection year, and map site (#s) separated by semicolons (Table 1).
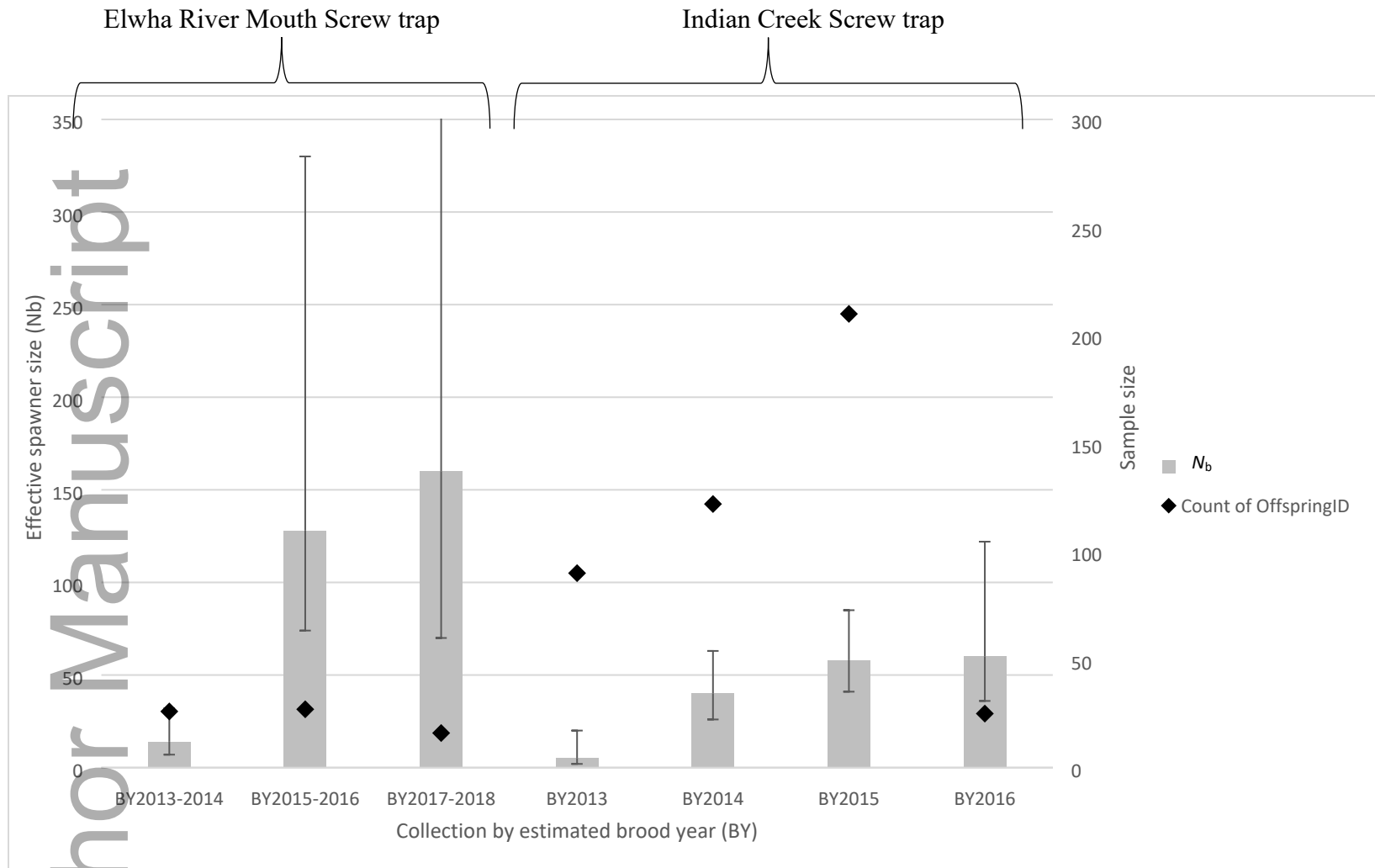
Figure 5. Effective spawner size estimates (N_b) with 95% confidence intervals and sample sizes (diamonds on secondary y-axis) for larval and juvenile screw trap collections of Pacific lamprey made in 2018 from the Elwha River basin. Collections are grouped by brood year as estimated by length-based ages in order from oldest to youngest brood years.
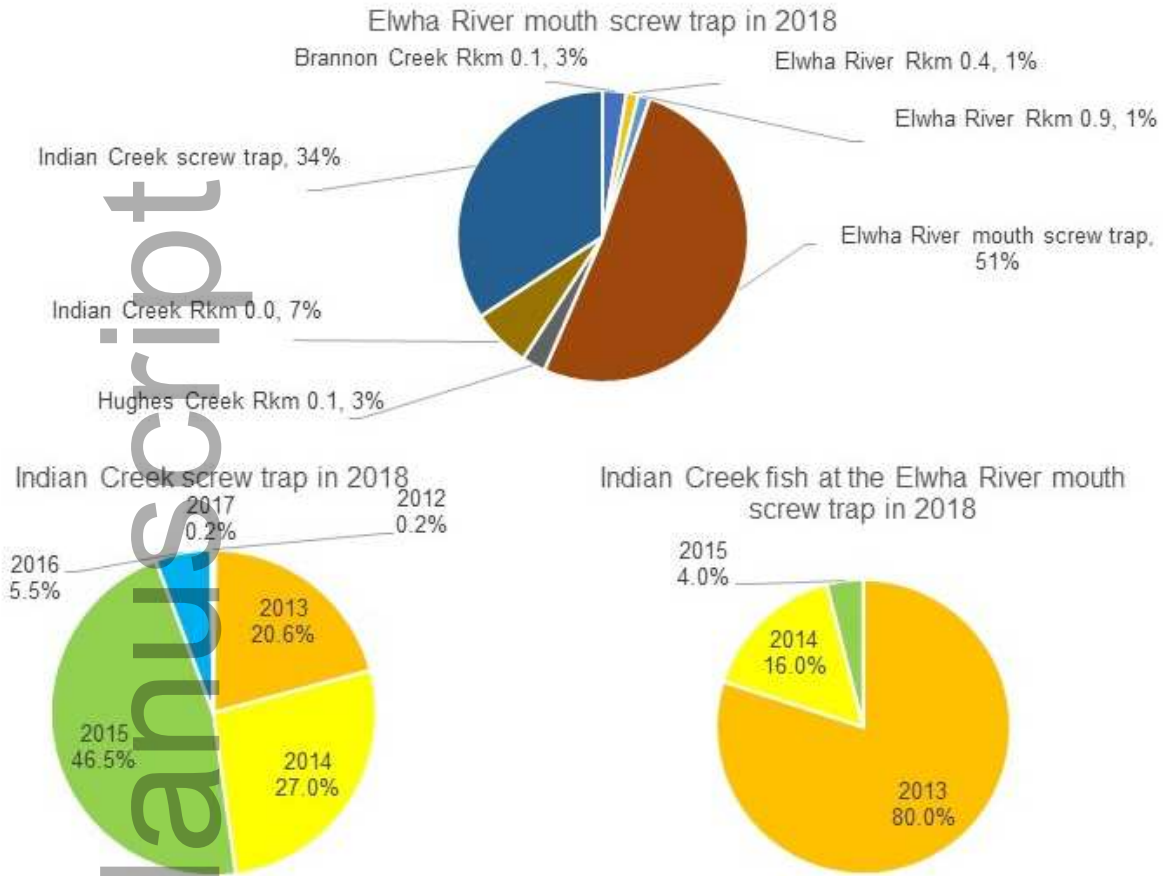
Figure 6. Relative percent composition of Pacific lamprey sibling-derived natal origins (listed by site) at the screw trap at the mouth of the Elwha River in 2018 (top), and length-based age composition of the Indian Creek screw trap (bottom left), and the Indian Creek fish "re-captured" at the Elwha River mouth (bottom right).

952

Author Manuscript