Are we underestimating the occurrence of sympatric populations?

Per Erik Jorde1 | Anastasia Andersson2 | Nils Ryman2 | Linda Laikre2

1Institute of Marine Research, His, Norway
2Division of Population Genetics, Department of Zoology, Stockholm University, Stockholm, Sweden

Correspondence
Linda Laikre, Division of Population Genetics, Department of Zoology, Stockholm University, Stockholm, Sweden.
Email: linda.laikre@popgen.su.se

Funding information
The Swedish Research Council Formas financed this work via grants to L.L.; financial support from the Swedish Agency for Marine and Water Management (L.L), the Swedish Research Council (N.R.) also aided in performing the work and the Carl Trygger Foundation (L.L).

Abstract
Sympatric populations are conspecific populations that coexist spatially. They are of interest in evolutionary biology by representing the potential first steps of sympatric speciation and are important to identify and monitor in conservation management. Reviewing the literature pertaining to sympatric populations, we find that most cases of sympatry appear coupled to phenotypic divergence, implying ease of detection. In comparison, phenotypically cryptic, sympatric populations seem rarely documented. We explore the statistical power for detecting population mixtures from genetic marker data, using commonly applied tests for heterozygote deficiency (i.e., Wahlund effect) and the STRUCTURE software, through computer simulations. We find that both tests are efficient at detecting population mixture only when genetic differentiation is high, sample size and number of genetic markers are reasonable and the sympatric populations happen to occur in similar proportions in the sample. We present an approximate expression based on these experimental factors for the lower limit of $F_{ST}$, beyond which power for STRUCTURE collapses and only the heterozygote-deficiency tests retain some, although low, power. The findings suggest that cases of cryptic sympatry may have passed unnoticed in population genetic screenings using number of loci typical of the pre-genomics era. Hence, cryptic sympatric populations may be more common than hitherto thought, and we urge more attention being diverted to their detection and characterization.

KEYWORDS
biodiversity monitoring, conservation management, genetic biodiversity, population genetic structure

1 | INTRODUCTION

Sympatric populations represent conspecific populations that coexist spatially during at least a part of their life cycle (Futuyama & Mayer, 1980; Mallet, Mayer, Nosil, & Feder, 2009). Such populations are of great interest in studies of ecological interaction and microevolutionary processes since their existence may represent the first steps of sympatric speciation processes (Maynard Smith, 1966; Via, 2001). They may reflect genetic adaptations to ecological niches and involve reproductive isolation occurring even in the absence of obvious migration barriers (Kawecki, 1996, 1997; Turelli, Barton, & Coyne, 2001).

From perspectives of management and conservation, sympatric populations are important to identify and monitor; they represent population diversity below the species level and such diversity has been documented to contribute to the portfolio effect in ecosystem stability (Schindler, Armstrong, & Reed, 2015; Schindler et al., 2010).
Sympatric populations have been described in a wide range of taxa and ecosystems. Marine examples include sympatric killer whale populations specializing on different diets (Ford et al., 1998) and blue whales (Attard, Beheregayar, & Möller, 2016) and beluga whales on summer foraging migration (Hauser, Laidre, Ruydom, & Richard, 2014). In coastal waters, genetically differentiated, sympatric populations or “ecotypes” have been described for the littorina snail (Ravinet et al., 2016) and the Atlantic cod (Barth et al., 2017; Knutsen et al., 2018). Sympatric “clades” have been described for the blue swimming crab (Ren et al., 2018) and diatom plankton (De Decker et al., 2018). Many examples are known from freshwater environments, including smelts, whitefish and salmonid fishes (reviewed by Taylor, 1999), and threespined sticklebacks (Marques et al., 2016). Salmonid fishes seem particularly well represented in this regard (brown trout: Allendorf, Ryman, Stennek, & Ståhl, 1976; Andersson, Jansson et al., 2017; Ryman, Allendorf, & Ståhl, 1979: whitefish: Bernatchez & Dodson, 1990; Østbye, Nesje, Bernatchez, Sandlund, & Hindar, 2005). Examples from terrestrial ecosystems are sympatric populations of the meadow butterfly (Ford, 1975, p. 78–108), Asian tiger mosquitoes (Guo et al., 2018) and ground beetles (Van Belleghem, De Wolf, & Hendrickx, 2016). A special case refers to sympatric “host races” of parasitic insects such as the apple maggot (Filchak, Rothelei, & Feder, 2000) and pea aphid (Peccoud, Ollivier, Plantegenest, & Simon, 2009). Host-specific races are also known for the brood-parasitic common cuckoo, although it is unclear whether these “gentes” represent different populations or genetic polymorphism within populations (Fossay et al., 2016). So-called “chromosome races” or “cytotypes” are known from small rodents that coexist in sympatry at least in zones of contact (house mouse: Corti & Rohlf, 2001; common scrub: Orlov et al., 2012). In plants, there is a large literature on co-occurring populations that differ in ploidy (Schönswetter et al., 2007). A common pattern in most, but not all, of these instances is that members of the sympatric populations differ to some extent in visual characteristics and this appears to have been a key feature for detecting such populations.

Sympatric populations may be described as cryptic when causal inspection had not previously revealed clear morphological or behavioural differences between them (Bickford et al., 2007). In such situations, the detection of sympatric populations typically requires some form of genetic data. Whether cryptic or not, there is a problem of demarcating sympatric populations against sympatric, closely related, sister species. Researchers adhering to a strict interpretation of the biological species concept may classify all sympatric, reproductively isolated populations as full species. There are thus likely to be differences among taxa, ecosystems and field of research in the detection of cryptic biodiversity and how this diversity is recognized at the species level or below (Bickford et al., 2007; Struck et al., 2018). There is also the problem of defining sympathy: At what spatial and temporal scales should coexistence be defined? Sympatric populations may coexist in the same area only relatively briefly, for example during seasonal feeding migration (beluga whales: Hauser et al., 2014) or during their entire lifespan (brown trout: Ryman et al., 1979; Palmé, Laikre, & Ryman, 2013). Sympathy is more readily defined within a confined environment such as a lake or an island than in the open ocean or in open terrestrial landscapes, and this may add to differences among taxa and environments in recognition of the existence and occurrence of sympathy.

We hypothesize that cryptic sympatric populations may have gone largely undetected and therefore might be under-reported in the literature. First, sympatric populations in general may be perceived as somewhat of an exception under the dominating ecological view emphasizing niche specialization and competitive exclusion (Harding, 1960), possibly diverting attention away from a systematic search for them. Hence, phenotypically cryptic, sympatric populations may go unnoticed except as chance detection in genetic screenings for, for example genetic diversity assessment.

Second, the statistical power for detecting sympatric populations may be relatively low, at least in the absence of obvious phenotypic differences. Without observable phenotypic differences and using genetic data alone, the classical test for the presence of more than one population in a sample of individuals relies on the Wahlund effect, that is, a deficiency of heterozygotes relative to the Hardy–Weinberg expectation (Rousset & Raymond, 1995; Waples, 2015). Deviations from Hardy–Weinberg genotype proportions may have gone unnoticed due to low power of detection (Fairbairn & Roff, 1980). Indirect evidence that power of genetic methods has been weak is the observation that most cases of reported sympathy appear to be coupled to phenotypic differences (Taylor, 1999; this study). Individuals can then be grouped according to phenotype, and potential genetic differences between groups are investigated. This kind of comparison is frequently associated with higher statistical power than a general exploration of Hardy–Weinberg deviations (Palmé et al., 2013).

Third, as microsatellites became the marker of choice over allozymes, it early became clear that technical artefacts (allelic dropout: Taberlet et al., 1996; short allele dominance: Wattler, Engel, Saumitou-Laprade, & Valero, 1998; stutter bands: Miller & Yuan, 1997) and segregating null alleles (Chapuis & Estoup, 2007) all could lead to deficiencies of heterozygotes unrelated to any population mixture (Band & Ron, 1997). Concerns over such artefacts may have led researchers to dismiss also real heterozygote deficiencies and thereby overlook signals from population mixtures in their samples (Waples, 2015). More generally, studies reporting heterozygote deficiencies often fail to follow up on those observations with further investigations, and this lack of follow-up investigations leaves the possibility of population mixture unresolved (Castric, Bernatchez, Belkhir, & Bonhomme, 2002).

Finally, statistical tools beyond the Hardy–Weinberg test have traditionally been lacking for detecting mixtures of phenotypically cryptic populations occurring in sympathy. Mixture of genetically differentiated populations leads not only to non-random association of alleles within loci, but also among alleles at different loci (so-called “linkage” disequilibrium or LD; Makela & Richardson, 1977) and potentially more powerful methods that explore both effects to detect population mixture were not generally available until the turn...
of the century (i.e., the \textit{structure} software: Pritchard, Stephens, & Donnelly, 2000). However, little is presently known about the statistical power of \textit{structure} relative to tests for heterozygote deficiency.

The purpose of the present paper is twofold. First, we review literature pertaining to sympatric populations. As pointed out above, there may be considerable differences among taxa and ecosystems with regard to how researchers recognize and interpret biological diversity. To maintain a level of consistency and uniformity, we therefore limit our review to freshwater salmonids, for which we have the most experience, sympathy is fairly easily defined and a relative rich literature exists. This review is pursued to summarize documentation of sympatric populations, particularly comparing the detection of cryptic, sympatric populations vs. non-cryptic detection, and further to find out whether commonly used genetic markers might have led to under-detection and hence under-reporting of sympatric populations. Second, we assess statistical power of detecting phenotypically cryptic populations from genetic data using computer simulations and focusing on realistic levels of genetic divergence, number of gene markers and sample sizes as revealed by the literature survey. The question addressed by these computer simulations is: what is the probability of detecting population admixture/structure from genotype data from a single sample or locality, that is, without additional information on habitat or phenotype differences?

2 \ MATERIALS AND METHODS

2.1 \ Literature survey

We carried out a literature survey on evidence for sympatric populations of salmonid fishes in freshwater environments using the Web of Science. We performed six topic searches using keyword combinations of “sympatric populations” AND either of the following “salmonid,” “trout,” “char,” “whitefish” OR “salmon.” The search included all years available in the database and was carried out in April 2018. In a next step, we examined the papers obtained for relevance with respect to our focus, that is, occurrence of sympatric populations in freshwater habitats. Further, we added nine papers that we knew of, but which did not appear in the searches. All in all, we included 80 studies in our survey. We classified the sympatric populations reported in these studies as cryptic if they were initially detected through genetics only, without prior identification of, or grouping based on, phenotypic, ecological or other divergence. The sympatric populations were classified as non-cryptic if the basis for detection was phenotypic differences and as ambiguous if they could not be classified as either cryptic or non-cryptic based on the information given in the studies.

2.2 \ Computer simulations

Simulations employing an in-house computer program were used to assess statistical power of detecting a significant indication of population mixture in genetic data from a sample of individuals when no phenotypic or non-genetic cues to population membership exist. Simulated sample data sets were generated by random sampling from two interconnected populations in approximate migration-drift equilibrium, and statistical tests included Hardy–Weinberg tests equilibration and tests derived from cluster analyses.

Each simulated population consisted of $N = 1,000$ diploid, sexually reproducing individuals with discrete (non-overlapping) generations. Populations were initiated with even sex ratios and with a number $L$ of freely combining (i.e., unlinked) loci with a specified allele frequency profile. Various numbers of alleles (2 or 20) and loci (up to 100) were used to represent popular genetic marker types and numbers commonly used in past and present population screenings as identified in our literature review (Table 1). In particular, we consider a set of 10 loci with (initially) 20 alleles each and refer to this set as the “microsatellite panel” and a set of 100 di-allelic loci referred to as the “SNP panel.” Smaller numbers of di-allelic loci were also simulated in order to represent allozyme-based studies. For the microsatellite panel, we incorporated mutations by randomly changing genes from their allelic state to one of the (19) other allelic states. We used a mutation rate of $u = 0.0005$, implying that one of the 2,000 genes in the population mutated per generation on average. The SNPs and allozymes were simulated without mutations. Simulations were initiated with even allele frequencies and were run for a sufficient number of generations (1,000) to thoroughly redistribute alleles within and among loci. Each generation after initiation (generation $t = 0$), $N$ haploid gametes, including $L$ loci plus the sex-determining locus, were drawn with replacement from males and from female parents, respectively, and merged into $N$ diploid offspring which immediately replaced the parental generation. Thus, generations were discrete (non-overlapping) and population size was kept exactly constant, while the sex ratio varied randomly (i.e., binomially with a mean of 0.5 and a standard deviation of 0.0158). Migration was simulated by exchanging a fixed number ($M$) of diploid individuals between the two populations each generation, following reproduction and mutation. A range of levels of genetic divergence between populations ($F_{ST}$) from 0.00025 to 0.39 was generated by exchanging different numbers of migrants ($M = 43, 23, 12, 5, 2.5, 1$ or $0$ per generation). Fractional numbers of migrants (e.g., 2.5) were accommodated by passing on the fractional part to the subsequent generation. Thus, in the case of $M = 2.5$, the actual number of migrants alternated between two and three in successive generations for an average of 2.5 per generation.

When sampling from the two populations, $n_1$ and $n_2$ diploid individuals were drawn from population 1 and 2, respectively, in generation $t = 1,000$ and both samples were pooled into a common file for statistical analyses. Different proportions of the two populations in samples were explored, from 1:1 (i.e., even representation) to 1:19 (highly skewed representation). The total sample size was set to cover the range over most empirical studies (Table 1), from 20 to 400 individuals combined ($n_1 + n_2$). When testing the case of no divergence (for assessment of alpha errors), that is $F_{ST} = 0$, a single, isolated population was simulated for 1,000 generations before the sample was drawn. Samples were drawn with replacement, in accordance with common—but typically not explicitly stated—assumptions of estimation procedures (Weir & Cockerham, 1984). The realized divergence ($F_{ST}$) between the two populations in the sample...
2.3 Statistical analyses of simulated data

Tests for heterozygote deficiency in the pooled samples (sample size = \( n_1 + n_2 \)) utilized the sampled genotypes, anonymized with respect to population of origin by erasing the population identifiers from the input file prior to statistical analyses. The calculations were carried out with *GENEPOP* option 1 (Hardy–Weinberg exact test) suboption 4 (global tests for heterozygote deficiency), with default dememorization number (10,000), number of batches (20) and iterations per batch (5,000). Results were summarized as the proportion of the 5,000 replicate simulation runs that yielded a significant, at the 5% level, global test.

The pooled and anonymized samples were further analysed for population structure with the command line version of the *STRUCTURE* software (v. 2.3: Falush, Stephens, & Pritchard, 2003; Pritchard, Xiaoquan, & Falush, 2010, sec- tion 5.1): Prob(\( K = 1 \)) = \( \exp \left[ \text{Ln Prob(data|} K = 1) \right] \) = \( \exp \left[ \text{Prob(data|} K = 1)] + \exp[\text{Prob(data|} K = 2)] + \exp[\text{Prob(data|} K = 3)] \right] \). Results of simulation runs were summarized as the proportion of replicate runs that yielded a Prob(\( K = 1 \)) less than 5% and interpreted as a significant (at the alpha = 5% level) detection of population mixing in the sample. In simulations involving a single population, that is with no true population mixing, the proportion of significant runs was interpreted as the alpha error of the test.

The discriminant analysis of the principal component method (DAPC: Jombart, Devillard, & Balloux, 2010), implemented in the R package *ADEGENET* (v.2.0.1: Jombart, 2008; Jombart & Ahmed, 2011), is also a potential tool for detecting the presence of individuals of different genetic origin within a sample. The *find.clusters* function calculates and reports the BIC (Bayesian information criterion) values for various number of clusters or populations (\( K \)) in the data. We did not find any application in the population genetics literature of using these BIC values to calculate posterior probabilities for the models, but the procedure is described in the general statistics literature on model selection (Burnham & Anderson, 2004, p. 275; Raftery, 1995) and is similar to that used for *STRUCTURE*. For testing the null hypothesis of a single population (\( K = 1 \)) in the sample, we calculated Prob(\( K = 1 \)) = \( \text{exp} \left(-1/2 \delta \text{BIC}_K = 1 \right) \text{sum} \left( \text{exp}(-1/2 \delta \text{BIC}_K = i) \right) \), where \( \delta \text{BIC}_K = i \) is the BIC value reported by the *find.clusters* function for \( K = 1 \) genetic clusters or populations, delta \( \text{BIC}_K = i \) is the difference between the BIC value for \( K = i \) and the lowest BIC value, and the summation is over \( i = 1 \ldots 10 \). We calculated the power and alpha errors of this test as the proportions of replicate runs that yielded Prob(\( K = 1 \)) < 0.05.

## 3 RESULTS

### 3.1 Literature survey

Review of the 80 papers identified in our literature survey revealed that for the case of salmonid fishes in freshwater habitats, sympatric populations have been reported in 136 cases in 135 localities in 17 countries, including at least 17 separate species (Table 2; Supporting Information Table S1). Arctic charr is the species with the largest number of reported sympatric cases with 39 localities where such existence has been documented. Sympatric populations have most commonly been found in freshwater lakes (108 cases), whereas river and creek habitats have been less commonly reported to harbour such populations (12 vs. 15 cases).

| Marker type | Number of loci | Number of individuals | \( F_{ST} \) |
|-------------|----------------|-----------------------|-------------|
|             | Average | Median | Min | Max | Average | Median | Min | Max | Average | Median | Min | Max |
| Allozymes   | 8.3     | 8      | 1   | 16  | 507     | 164    | 48  | 6159 | 0.065  | 0.050  | 0.010 | 0.200 |
| Microsatellites | 8.7    | 6      | 5   | 22  | 139     | 81     | 22  | 636  | 0.106  | 0.062  | 0.007 | 0.381 |
| SNPs        | 1,092.4 | 94     | 94  | 3,093 | 134    | 48     | 24  | 744  | 0.071  | 0.036  | 0.010 | 0.280 |

Note: Min: the smallest number of loci/individuals used and the smallest \( F_{ST} \) observed; Max: the largest number of loci/individuals employed and the largest \( F_{ST} \) observed.
Based on the classification described in the methods section, we find that non-cryptic sympatric populations are much more commonly reported than cryptic ones. Non-cryptic populations have been documented in 98 localities involving 17 separate species (Table 2), whereas cryptic populations have only been documented in nine freshwater localities (seven lakes and two rivers) and for five species: Arctic charr, brown trout, lake trout, Atlantic salmon and Chinook salmon (Aykanat et al., 2015; Marin, Coon, Carson, Debes, JORDE ET AL. | 4015

| Table 2  | Number of cases of sympatric populations in separate salmonid fishes in freshwater habitats classified as cryptic, non-cryptic and ambiguous (see text and Supporting Information Table S1 for details) |
|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Species                          | Cryptic | Non-cryptic | Ambiguous | All sympatic |
| Arctic char (Salvelinus alpinus) | 4        | 25          | 11         | 40          |
| Arctic char species complex       | 0        | 10          | 2          | 12          |
| Atlantic salmon (Salmo salar)     | 1        | 2           | 0          | 3           |
| Brook char (Salvelinus fontinalis) | 0        | 1           | 16         | 17          |
| Brown trout (Salmo trutta)        | 2        | 3           | 0          | 5           |
| Chionook salmon                   | 1        | 1           | 0          | 2           |
| Cisco (Coregonus albula; Coregonus fontane) | 0        | 2           | 0          | 2           |
| European whitefish (Coregonus lavaretus) | 0        | 14          | 0          | 14          |
| Lake trout (Salvelinus namaycush) | 1        | 1           | 0          | 2           |
| Lake whitefish                    | 0        | 9           | 0          | 9           |
| Mediterranean/Fibreni trout       | 0        | 1           | 0          | 1           |
| Ohrid trout (Salmo letrica typicus; Salmo letrica aestivalis) | 0        | 1           | 0          | 1           |
| Pygmy whitefish (Prosopium coulteri) | 0        | 1           | 0          | 1           |
| Rainbow smelt (Osmerus mordax)    | 0        | 1           | 0          | 1           |
| Rainbow/Steelhead trout           | 0        | 5           | 0          | 5           |
| Sockeye/kokanee salmon            | 0        | 11          | 0          | 11          |
| Whitefish (Coregonus spp.)        | 0        | 8           | 0          | 8           |
| Total (all species)               | 9        | 98          | 29         | 135         |

*Loch Awe in Scotland houses both two cryptic Arctic charr and two non-cryptic brown trout populations and is included both in the total number localities for both cryptic (9) and non-cryptic (98) populations. Thus, the total number of localities with sympatric populations identified in this study is 135.
Differences in resource use have been reported in several cases of non-cryptic sympatry. In lake habitats, such differences include food niches (21 lakes), spawning time (3 lakes), spawning place (4), anadromous vs. resident strategy (2), both spawning time and place differences (3), and both spawning time and habitat differences (2).

All non-cryptic sympatry in creeks is associated with anadromous vs. resident life history strategy (11 cases). In rivers, such differences are also found (three rivers) but here spawning time differences is another diverging factor (three rivers), whereas food niche separation has not been reported in creeks or rivers (Supporting Information Table S1).

In the cases of cryptic sympatry, clear life history strategy differences have only been observed in the case of Atlantic salmon in the Teno River (Aykakat et al., 2015). There are some indications of trophic divergence in sympatric charr in Lochs Maree and Stack (Adams, Wilson, & Ferguson, 2008), and extensive screening of the two cases of cryptic brown trout populations in tiny mountain lakes in Sweden found no trophic divergence but small growth and maturation differences as well as a tendency for a spacial separation at spawning (Andersson, Johansson, Sundborn, Ryman, & Laikre, 2017; Palmé et al., 2013; Ryman et al., 1979). Growth differences between cryptic, sympatric populations have been reported in a total of six cases (five lakes and one river).

Microsatellites and/or allozymes were the most frequently used markers for investigating genetic structure and had been employed in 40 vs. 21 studies, respectively. Typically, 1–16 loci were employed for allozymes and 5–22 for microsatellites. Four studies had used SNPs, five employed gene sequencing and several studies used combinations of different markers (Supporting Information Table S1). Studies identifying cryptic sympatric populations were based exclusively on heterozygote deficiency in one case (no heterozygotes observed; Ryman et al., 1979) and exclusively on STRUCTURE software in two cases (Aykakat et al., 2015; Marin et al., 2016). Two studies used a combination of heterozygote-deficiency tests and the STRUCTURE software (Palmé et al., 2013; Wilson et al., 2004), and one study applied an assignment software exclusively (Oncor; Smith & Engle, 2011).

We wanted to find out if studies reporting cryptic vs. non-cryptic sympatric populations differed with respect to sample size, number of loci or degree of genetic divergence. For such a comparison, we selected studies reporting all the relevant quantities, that is, number of fish, number of loci and significant \( F_{ST} \) (or equivalent). We limited our selection to studies using allozymes and/or microsatellites, since these were the most frequently applied markers. Of the 80 studies, 35 fulfilled these criteria and they represent 58 localities, seven with cryptic populations and 51 with non-cryptic ones (Supporting Information Tables S2). \( F_{ST} \) was consistently higher among cryptic populations as compared to non-cryptic ones using allozymes, microsatellites or a combination of both (Table 3). However, statistical significance was only obtained for allozymes using a \( t \) test (median test non-significant). Similarly, a larger number of individuals had been sampled in studies reporting cryptic populations based on allozymes or both markers as compared to studies reporting non-cryptic populations. However, this difference was only significant for the median test. The number of loci appointed were essentially the same (Table 3).

### 3.2 Computer simulations

The overall impression from the computer simulations (Supporting Information Table S3) is that STRUCTURE was superior to DAPC and also more powerful than the heterozygote-deficiency test to detect...
The power of STRUCTURE to detect population mixtures fell rather rapidly with declining levels of genetic divergence, and the rate of this decline was dependent on sample size (Figure 2) and number of independent genes (i.e., observations) per sampled individual. At the same time, nearly all (>99%) SNPs were still polymorphic (two alleles each), and samples with this marker therefore had a similar number of 100 × (2 − 1) = 100 independent genes per individual sampled. Statistical power of the two marker panels was therefore, as expected, fairly similar (cf. Figure 1).

The power of STRUCTURE to detect population mixtures fell rather rapidly with declining levels of genetic divergence, and the rate of this decline was dependent on sample size (Figure 2) and number of loci (Figure 3). For a given sample size and genetic marker panel, a reduction in $F_{ST}$ by a factor of 2 sometimes reduced power from very high to very low levels. For example, using the microsatellite panel and a sample size of $n = 100$ individuals (green line in Figure 2, upper right), power was 90% to detect an even mix of populations that diverged by $F_{ST} = 0.0435$ but power declined to only 17.5% when divergence was $F_{ST} = 0.0250$ (Supporting Information Table S3). For a similar sample size but using the SNP panel, the major drop in power (from 77.6% to just 2.3%) occurred at a somewhat lower $F_{ST}$ level, between 0.025 and 0.010 (cf. Figure 2, lower panel and Supporting Information Table S3). These major drops in statistical power indicate the existence of certain “threshold” levels of divergence necessary for the STRUCTURE test to effectively detect population mixture. The effect of sample size on power for this test was proportional and predictable (Figure 2, right panels): Expressed in terms of this threshold $F_{ST}$, a reduction in sample size by a given factor led to a corresponding increase in the $F_{ST}$ level that was required to maintain statistical power of detection. The effect on power of increasing or decreasing number of loci, assuming the same number of alleles per locus, was similar to the effect of increasing or decreasing sample size (Figure 3): An increase in number of loci typically resulted in a proportional reduction in the threshold $F_{ST}$.

For the heterozygote-deficiency test, power declined more slowly with reduced $F_{ST}$ and eventually overtook as the more powerful test when $F_{ST}$ fell below the threshold level for STRUCTURE (Figure 1). Nevertheless, power to detect heterozygote deficiencies at these low levels of divergence was poor in absolute terms, typically below 0.2, and implying that mixtures of weakly differentiated populations are likely to go unnoticed with either method, even if using large sample sizes ($n = 400$ individuals: Figure 2). In this parameter region with overall low power, there was generally a poor relationship between the realized $F_{ST}$ in the individual computer runs and the outcome of the statistical tests in those runs (Figure 4, upper panels). This lack of correspondence between level of...
divergence and test outcome likely reflected an increase in alpha errors (i.e., type I errors) relative to true rejections when true divergence was low. This sometimes resulted in contrasting outcomes of the two tests, with some runs being significant for the heterozygote-deficiency test and others for the STRUCTURE test and relatively few with both tests being significant. In the parameter region of higher overall power (Figure 4, bottom panels), such differences among tests rarely occurred and runs that were significant for the least powerful test (here, the heterozygote-deficiency test) were almost always significant also for the more powerful one (STRUCTURE).

With unequal sampling of the two populations, power of detecting mixtures declined with increasing skewness in population representation in the sample (Figure 5). While mild deviations of 30/70 from even proportions (50/50) only resulted in a minor effect on power, highly skewed representations of the two populations in the sample (10/90 and 5/95) typically limited power to a substantial

![Graphs showing power of detecting significant evidence for different size (n individuals), genotyped for the microsatellite panel (top) and the SNP panel (bottom) and tested with the GENEPOP heterozygote-deficiency test (left) and STRUCTURE test for K > 1 (right). Dots on the left margins indicate proportions of significant runs from a single, panmictic population and represent the alpha errors of the tests (note that some dots overlap).]
extent (Figure 5). STRUCTURE was somewhat more affected by uneven representation than was the heterozygosity-deficiency test but for highly divergent populations ($F_{ST} > 0.1$) power remained reasonable high ($> 0.5$) for both methods also with highly skewed representation (5/95 proportions).

Alpha errors, that is, the proportion of simulation runs with a single, panmictic population only but which nevertheless resulted in a significant test outcome for mixture, were always close to the nominal alpha (5%) level for the heterozygote-deficiency test (cf. Figure 2), ranging between 0.042 and 0.061 over simulations representing a wide range of sample sizes (Supporting Information Table S3). The STRUCTURE test also tended to display alpha errors in the vicinity of the alpha level, but was more variable and ranged between 0.01 and 0.16 (Figure 2; Supporting Information Table S3).

4 | DISCUSSION

In our literature case study of sympatric populations using salmonid fishes in freshwater habitats as models, we found that the majority of reported cases—98 out of 136—refer to non-cryptic populations that were identified by differences in phenotypic and/or behavioural traits. Only nine of the 136 examples that we found refer to cryptic, sympatric populations, leaving the impression that such populations are rare. Moreover, we found that genetic divergence was on average higher between cryptic than between non-cryptic populations. This is contrary to expectation because phenotypically cryptic populations are commonly thought to be evolutionary young (see review and discussion by Fišer, Robinson, & Malard, 2018) and therefore less differentiated at neutral loci. Thus, the finding that cryptic populations instead tended to be more differentiated suggests that reported cases provide a biased view and represent situations where statistical power was high. Ecological divergence in sympatric populations appear to differ with respect to habitat but in the case of cryptic sympathy, obvious genetic differences are typically associated with only weak and unclear resource use divergence leaving the evolutionary mechanisms behind such structuring presently unclear.

Recent works that were not included in our literature review find refined food niche separation in three sympatric genetically divergent groups of brown trout in Loch Leidon, Scotland (Piggott et al., 2018), and genomic signals indicating selection between the non-cryptic life history forms of brown trout of Loch Maree, Scotland (Jacobs, Hughes, Robinson, Adams, & Elmer, 2018). Evidence for sympatric genetic divergence between non-cryptic Arctic charr populations was reported by Salisbury et al. (2018) in Ramah Lake in Labrador, Canada, and Guðbrandsson et al. (2018) found gene expression divergence during early development among non-cryptic populations of this species in Lake Thingvallavatn on Iceland.

We used computer simulations to evaluate the statistical power of methods that utilize genetic markers for detecting sympatric populations in a sample of individuals, without prior groupings. Among methods, the Hardy–Weinberg test represents the classical approach and different variants of this test have been developed. For the particular purpose of detecting Wahlund effects, the exact heterozygote-deficiency test (Rousset & Raymond, 1995) was used as this seems the most appropriate and has seen wide use in empirical studies (the original paper was cited >600 times at Web of Science by June 2018, but most papers using this method probably cite the GENEPOP papers, Raymond & Rousset, 1995; Rousset, 2008; instead, with a combined
15,000 citations). Statistical power for Hardy–Weinberg tests in general and deficiency of heterozygotes in particular has been evaluated in the earlier literature for a variety of applications, including detection of cryptic population structure (Chakraborty & Zhong, 1994; Salanti, Amountza, Ntzani, & Ioannidis, 2005). Methods that simultaneously utilize LD and heterozygote deficiencies have been developed, most notably in the STRUCTURE software (Pritchard et al., 2000), and been widely applied (>16,000 citations). Simulation studies characterizing statistical properties of STRUCTURE include Castric et al. (2002), Manel, Berthier, and Luikart (2002), Evanno, Regnaut, and Goudet (2005), Latch, Dharmarajan, Glaubitz, and Rhodes (2006), Patterson, Price, and Reich (2006), Waples and Gaggiotti (2006), Anderson and Dunham (2008), Frantz, Cellina, Krier, Schley, and Burke (2009), Schwartz and McKelvey (2009), Jombart et al. (2010), Kalinowski (2011), Aurelle and Ledoux (2013), Neophytou (2014), Puechmaille (2016), Janes et al. (2017) and Wang (2017). These studies mainly addressed the problem of correctly estimating the number \( K \) of populations represented by a set of samples, and only Castric et al. (2002) specifically addressed the related but different problem of detecting a significant signal for \( K > 1 \). Nevertheless, these studies have yielded relevant information on the question at hand by demonstrating that STRUCTURE has problems with detecting populations with low genetic divergence \( F_{ST} \) below about 0.01–0.05 (Castric et al., 2002; Latch et al., 2006; Waples & Gaggiotti, 2006). More specifically, Patterson et al. (2006) conjectured that any genetic clustering method, including STRUCTURE, will fail when \( F_{ST} < 1/\sqrt{D} \), where \( D \) is the data size or the number of independent genes surveyed. It is not clear exactly when STRUCTURE “fails” in the present context of power analysis, but if we take a power of 10% as the point where there is no longer any obvious signal of population mixture, the “threshold” \( F_{ST} \) values that correspond to this power in the case of the SNP panel are about 0.027, 0.014, 0.011, 0.007 and 0.005 for sample sizes \( n = 20, 50, 100, 200 \) and 400, respectively (cf. Figure 2, bottom right). Because these are biallelic loci, there is only one independent allele per locus, so that with 100 loci \( 1/\sqrt{D} = 1/\sqrt{n \times 100 \times (2-1)} \) = 0.022, 0.014, 0.010, 0.007 and 0.005, in good agreement with the theoretical predictions. The agreement is less good for the microsatellite panel (with corresponding \( F_{ST} \) values of about 0.050, 0.030, 0.015,
0.011 and 0.006, respectively; cf. Figure 2, top right), implying somewhat lower power for them, probably because alleles at the same locus are not entirely independent and that more alleles occurred in low frequencies as compared to the di-allelic SNPs. The expression is nevertheless useful also for microsatellites as a rough guideline for conditions where the statistical power of the STRUCTURE test becomes too low to be of practical use. However, this guideline refers to population mixtures in equal proportions and our simulation results show that power of STRUCTURE is reduced when populations are unequally represented in the sample (Figure 5, right), as also found earlier (Puechmaille, 2016; Wang, 2017). This reduction in power and subsequent increase in threshold $F_{ST}$ appears to be directly related to the relative proportion of the two populations in the sample, or $r = n_1/n_2$ (where $n_2$ is the larger, so that $r \leq 1$). Using $r$, we may tentatively consider a modified expression for the threshold $F_{ST}$ for STRUCTURE:

$$1/\sqrt{D} = 1/\sqrt{n \times r \times L \times (a - 1)}$$

(1)

where $n$ is the sample size, $L$ the number of loci and $a$ number of alleles per locus. With $r = 5/95, 10/90, 30/70$ and $50/50$, as in our simulations (Figure 5), we then obtain for the 100 locus SNP panel...
the following threshold $F_{ST}$ values: 0.010, 0.015, 0.030 and 0.044, respectively. In comparing these numbers with simulation results, we extract $F_{ST}$ values representing a power of 10% from Figure 5, bottom right, which turns out to be 0.011, 0.012, 0.029 and 0.053, respectively, in reasonable agreement with the calculations. These numbers imply that, with a skewed population representation, say 5/95, the level of genetic differentiation ($F_{ST}$) needs to be much higher, here 0.053/0.011 = 4.8 times higher, in order to achieve the same power of detection as with even population representation.

In comparison with results for STRUCTURE, power for the heterozygote-deficiency test decreased more gradually with decreasing $F_{ST}$ (cf. Figures 1 and 2) and the concept of a lower “threshold” $F_{ST}$ value is even less clear. Nevertheless, this test generally retained somewhat higher power than the STRUCTURE test at low levels of divergence, implying that a signal of sympatric mixture is slightly more likely to be picked up with the heterozygote-deficiency test. Apart from this slight advantage at low levels of divergence, the heterozygote-deficiency test displayed similar patterns of power depending on sample size, number of loci and population representation in the samples, as for STRUCTURE.

As discussed in the introduction, there are a number of circumstances that may have conspired against detecting sympatric populations in empirical genetic studies. Apart from situations where a significant test outcome is being dismissed or misinterpreted, our findings add insights into non-significant outcomes and aid in interpreting the current literature pertaining to the occurrence of cryptic, sympatric populations. Briefly, computer simulations verified that weakly divergent populations are likely to go undetected regardless of statistical test employed when number of loci and/or sample sizes are small. While we do not know what level of divergence may characterize real cryptic, sympatric populations, the literature review of non-cryptic populations showed that these diverged with a broad range of $F_{ST}$ from 0.003 to 0.497, with an average of 0.095.

While recent advances in genomics have hugely expanded the number of loci that can be genotyped, most studies until recently were severely limited by available protein staining technology (for allozymes) or species-specific primers (for microsatellites). For studies employing microsatellites, numbers of loci in our literature survey ranged from 5 to 25 with a mean of 10, and for allozymes, the number of (polymorphic) loci was even lower. Combined with a moderate sample size, we conclude that most studies would not have been able to detect sympatric populations from genetic data alone. In the present genomics era increasing the number of loci substantially is brought attention to the phenomenon as not just a technical artefact with a few samples.

A complicating factor relating to detection of cryptic sympathy that we have not addressed here is that degree of divergence most likely differs in different regions of the genome. Such differences might explain the difficulty in detecting the two cryptic populations of brown trout in Lakes Trollsvattnet that we have reported and monitored over time (Andersson, Jansson et al., 2017; Andersson, Johansson et al., 2017; Jorde & Ryman, 1996; Palmé et al., 2013) with six microsatellites as compared to 14 allozyme loci. In fact, the degree of divergence between these population using allozymes is estimated as $F_{ST} = 0.1$, whereas when applying ~3,000 SNPs, we find a lower $F_{ST} = 0.03$ (Andersson, Jansson et al., 2017). Clearly, more research is needed into the issue of cryptic sympathy to understand the evolutionary background to their existence. From the perspective of conservation management, mapping the existence of this type of biodiversity over space and monitoring such existences over time is important.

ACKNOWLEDGEMENTS

We thank three anonymous reviewers and the subject editor for important comments on a previous version of this manuscript.

DATA ACCESSIBILITY

Results of the computer simulations and R-scripts for generating the figures described herein are available on Dryad https://doi.org/10.5061/dryad.68r9h2.

AUTHOR CONTRIBUTION

L.L., N.R., A.A. and P.E.J. designed the study; A.A. did the literature survey; and P.E.J. performed the computer simulations and wrote the first draft. All the authors contributed to the final version.

ORCID

Per Erik Jorde http://orcid.org/0000-0001-5515-7257
Anastasia Andersson http://orcid.org/0000-0002-5698-4948
Nils Ryman http://orcid.org/0000-0003-3342-8479
Linda Laikre http://orcid.org/0000-0001-9286-3361

REFERENCES

Adams, C. E., Wilson, A. J., & Ferguson, M. M. (2008). Parallel divergence of sympatric genetic and body size forms of Arctic char, Salvelinus alpinus, from two Scottish lakes. *Biological Journal of the Linnean*
Knutsen, H., Jorde, P. E., Hutchings, G. A., Hemmer-Hansen, J., Grønkjør, E. K., Dharmarajan, G., Glaubitz, J. C., & Rhodes, O. E. Jr (2006). Kawecki, T. J. (1996). Sympatric speciation driven by beneficial mutations. Proceedings of the Royal Society B: Biological Sciences, 263, 1515–1520. https://doi.org/10.1098/rspb.1996.0221

Kawecki, T. J. (1997). Sympatric speciation via habitat specialization driven by deleterious mutations. Evolution, 51, 1751–1763. https://doi.org/10.1111/1558-5646.1997.tb05099.x

Knutsen, H., Jorde, P. E., Hutchings, G. A., Hemmer-Hansen, J., Granrøe, P., Mose Jørgensen, K.-E., Olsen, E. M. (2018). Stable coexistence of genetically divergent Atlantic cod ecotypes at multiple spatial scales. Evolutionary Applications, 1–13. https://doi.org/10.1111/eva.12640

Latch, E. K., Dharmarajan, G., Glaubitz, J. C., & Rhodes, O. E. Jr (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conservation Genetics, 7, 295–302. https://doi.org/10.1007/s10592-005-9098-1

Makela, M. E., & Richardson, R. H. (1977). The detection of sympatric sibling species using genetic correlation analysis. I. Two loci, two gamodemes. Genetics, 86, 665–678.

Mallet, J., Mayer, A., Nosil, P., & Feder, J. L. (2009). Space, sympathy and speciation. Journal of Evolutionary Biology, 22, 2332–2341. https://doi.org/10.1111/j.1420-9101.2009.01816.x

Manel, S., Berthier, P., & Luikart, G. (2002). Detecting wildlife poaching: Identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conservation Biology, 16, 650–659. https://doi.org/10.1046/j.1523-1739.2002.00576.x

Marin, K., Coon, A., Carson, R., Debes, P. V., & Fraser, D. J. (2016). Striking phenotypic variation yet low genetic differentiation in sympatric lake trout (Salvelinus namaycush). PLoS One, 11, e0162325. https://doi.org/10.1371/journal.pone.0162325

Marques, D. A., Lucek, K., Meier, J. I., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2016). Genomics of rapid incipient speciation in sympatric threespine stickleback. PLoS Genetics, 12(2), e1005887.

Maynard Smith, J. (1966). Sympatric speciation. American Naturalist, 100, 637–650. https://doi.org/10.1086/282457

Miller, M. J., & Yuan, B.-Z. (1997). Semiautomated resolution of overlapping stutter patterns in genomic microsatellite analysis. Analytical Biochemistry, 251, 50–56. https://doi.org/10.1006/abio.1997.2234

Neophytou, C. (2014). Bayesian clustering analyses for genetic assignment and study of hybridization in oaks: Effects of asymmetric phylogenies and asymmetric sampling schemes. Tree Genetics & Genomes, 10, 273–285. https://doi.org/10.1007/s11295-013-0680-2

Orlov, V. N., Borisov, Y. M., Cherepanova, E. V., Grigor’eva, O. O., Shesmak, A. G., & Sycheva, V. B. (2012). Narrow hybrid zone between Moscow and Western Dvina chromosomal races and specific features of population isolation in common shrew Sorex araneus (Mammalia). Russian Journal of Genetics, 48, 70–78. https://doi.org/10.1134/S1022795412010152

Østbye, K., Nasje, T. F., Bernatchez, L., Sandlund, O. T., & Hindar, K. (2005). Morphological divergence and origin of sympatric populations of European whitefish (Coregonus lavaretus L) in Lake Fenem, Norway. Journal of Evolutionary Biology, 18, 683–702.

Palme, A., Laike, L., & Ryman, N. (2013). Monitoring reveals two genetically distinct brown trout populations remaining in stable sympatry over 20 years in tiny mountain lakes. Conservation Genetics, 14, 795–808. https://doi.org/10.1007/s10592-013-0475-x

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenvector analysis. PLoS Genetics, 2, 2074–2093.

Peccoud, J., Ollivier, A., Plantegenest, M., & Simon, J.-C. (2009). A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. PNAS, 106, 7495–7500. https://doi.org/10.1073/pnas.0811117106

Piggott, C. V. H., Verspoor, E., Greer, R., Hooker, O., Newton, J., & Adams, C. E. (2018). Phenotypic and resource use partitioning amongst sympatric, lacustrine brown trout, Salmo trutta. Biological Journal of the Linnean Society, 124, 200–212. https://doi.org/10.1093/biolinnean/bly032

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics, 155, 945–959.

Pritchard, J. K., Xiaoquan, W., & Falush, D. (2010). Documentation for structure software: Version 2.3. Retrieved from: https://web.stanford.edu/group/pritchardlab/software/release_versions/v2.3.4/structure_doc.pdf

Puechmaillie, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. Molecular Ecology Resources, 16, 608–627. https://doi.org/10.1111/1755-0998.12512

Raferty, A. E. (1995). Bayesian model selection in social research. Sociological Methodology, 25, 11–163.

Ravinet, M., Westram, A., Johannesson, K., Butlin, R., André, C., & Panova, M. (2016). Shared and nonshared genomic divergence in parallel ecotypes of Littorina saxatilis at a local scale. Molecular Ecology, 25, 287–305. https://doi.org/10.1111/mec.13332

Raymond, M., & Rouset, F. (1995). Genepop (version 1.2): Population genetics software for exact tests and ecumenism. The Journal of Heredity, 86, 248–249. https://doi.org/10.1093/oxfordjournals.jhered.a111573

Ren, G., Miao, G., Ma, C., Lu, J., Yang, X., & Ma, H. (2018). Genetic structure and historical demography of the blue swimming crab (Portunus pelagicus) from southeastern sea of China based on mitochondrial COI gene. Mitochondrial DNA Part A, 29, 192–198. https://doi.org/10.1080/24701394.2016.1261855

Rousset, F. (2008). Genepop’007: A complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources, 8, 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Rousset, F. (2013). Genetic structure and selection in subdivided populations. Princeton, NJ: Princeton University Press.

Rousset, F., & Raymond, M. (1995). Testing heterozygote excess and deficiency. Genetics, 140, 1413–1419.

Ryman, N., Allendorf, F. W., & Ståhl, G. (1979). Reproductive isolation with little genetic divergence in sympatric populations of brown trout (Salmo trutta). Genetics, 92, 247–262.

Salanti, G., Amontza, G., Ntzani, E. E., & Ioannidis, J. P. A. (2005). Hardy-Weinberg equilibrium in genetic association studies: An empirical evaluation of reporting, deviations, and power. European Journal of Human Genetics, 13, 840–848. https://doi.org/10.1038/sj.ejhg.5201410
Salisbury, S. J., Booker, C., McCracken, G. R., Knight, T., Keefe, D., Perry, R., & Ruzzante, D. E. (2018). Genetic divergence among and within Arctic char (Salvelinus alpinus) populations inhabiting landlocked and sea-accessible sites in Labrador, Canada. Canadian Journal of Fisheries and Aquatic Sciences, 75, 1256–1269. https://doi.org/10.1139/cjfas-2017-0163

Schindler, D. E., Armstrong, J. B., & Reed, T. E. (2015). The portfolio concept in ecology and evolution. Frontiers in Ecology and Environment, 13, 257–263. https://doi.org/10.1890/140275

Schindler, D. E., Hilborn, R., Chasco, B., Boattright, C. P., Quinn, T. P., Rogers, L. A., & Webster, M. S. (2010). Population diversity and the portfolio effect in an exploited species. Nature, 465, 609–612. https://doi.org/10.1038/nature09060

Schönswetter, P., Lachmayer, M., Lennert, C., Prehsler, P., Rechnitzer, S., Reich, D. S., ... Suda, J. (2007). Sympatric diploid and hexaploid cytotypes of Senecio carniolicus (Asteraceae) in the Eastern Alps are separated along an altitudinal gradient. Journal of Plant Research, 120, 721–725. https://doi.org/10.1007/s10265-007-0108-x

Schwartz, M. K., & McKelvey, K. S. (2009). Why sampling scheme matters: The effect of sampling scheme on landscape genetic results. Conservation Genetics, 10, 441–452. https://doi.org/10.1007/s10592-008-9622-1

Smith, C. T., & Engle, R. (2011). Persistent reproductive isolation between sympatric lineages of fall Chinook salmon in White Salmon River, Washington. Transactions of the American Fisheries Society, 140, 699–715. https://doi.org/10.1080/00028487.2011.584490

Struck, T. H., Feder, J. L., Bendiksky, M., Birkeland, S., Cerca, J., Gusarov, V. L., ... Dimitrov, D. (2018). Finding evolutionary processes hidden in cryptic species. Trends in Ecology and Evolution, 33, 153–163. https://doi.org/10.1016/j.tree.2017.11.007

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Gracilarians gracilis (Rhodophyta). Molecular Ecology, 7, 1569–1573. https://doi.org/10.1046/j.1365-294x.1998.00477.x

Wattier, R., Engel, C. R., Saumitou-Laprade, P., & Valera, M. (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite loci: Experimental evidence at the dinucleotide locus Gv1CT in Gracilaria gracilis (Rhodophyta). Molecular Ecology, 7, 1569–1573. https://doi.org/10.1046/j.1365-294x.1998.00477.x

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358–1370.

Wilson, A. J., Gíslason, D., Snorrason, S. S., Adams, C. E., Alexander, G., ... Ferguson, M. M. (2004). Population genetic structure of Arctic Charr, Salvelinus alpinus from northwest Europe on large and small spatial scales. Molecular Ecology, 13, 1129–1142. https://doi.org/10.1111/j.1365-294x.2004.02149.x

How to cite this article: Jorde PE, Andersson A, Ryman N, Laikre L. Are we underestimating the occurrence of sympatric populations? Mol Ecol. 2018;27:4011–4025. https://doi.org/10.1111/mec.14846

Additional supporting information may be found online in the Supporting Information section at the end of the article.