ICES Journal of Marine Science

ICES Journal of Marine Science (2020), 77(5), 1658–1665. doi:10.1093/icesjms/fsaa061

Food for Thought

Serendipity and me

Robin S. Waples [1]*

[1]NOAA Fisheries, Northwest Fisheries Science Center, 2725 Montlake Blvd East, Seattle, WA, 98112, USA

*Corresponding author: tel: + 1 206 860 3254; e-mail: robin.waples@noaa.gov.

Waples, R. S. Serendipity and me. – ICES Journal of Marine Science, 77: 1658–1665.

Received 24 February 2020; revised 16 March 2020; accepted 17 March 2020; advance access publication 21 May 2020.

How much of your scientific career has unfolded as you planned, and how much has been shaped by blind luck? I suspect the latter has been more important than many of us realize, or at least acknowledge, but as Louis Pasteur said, "Chance favors only the prepared mind"—implying that we have at least some control over how random events affect our lives. Here, I discuss how large and small chance events have affected the trajectory of my scientific career.

Keywords: allozymes, chance, cryptic species, effective population size, scientific career

Introduction

What if one could adapt Steven Jay Gould's idea and play your life's tape over again—what would be different, and what would be the same? It is tempting to think of the long arc of one's career as being the result of foresight, persistence, and hard work. In reality, however, random events large and small shape our lives more than we think, or at least more than we generally acknowledge. Below I recount several instances where serendipity has intervened to affect my scientific career.

What should I do when I grow up?

I did not have a good answer to that question for a long time. Unlike Charles Darwin or E. O. Wilson, I was not a passionate naturalist as a kid. Growing up in Iowa, I spent a lot of time outdoors, but mostly to dig a tunnel, go swimming or ride my bike, have a snowball fight, find a mud puddle to wrestle in, or engage in politically incorrect activities involving cap pistols. This lack of a clear career trajectory lasted through college, where I majored in American Studies because it allowed me to take a smorgasbord of fascinating courses from people like Margaret Mead (who came up once a week from New York), Charles Reich (his Law School course evolved into *The Greening of America*), Vincent Scully (legendary lecturer and critic of American architecture), and Erich Segal (classics professor who wrote *Love Story* in his spare time). If I had been paying attention, I could have taken a course in something called "ecology" by G. Evelyn Hutchinson, who practically invented the field—but I was not interested in science at the time.

After college, my first real job was at Punahou School in Honolulu, teaching English and coaching water polo. The players

Food for Thought articles are essays in which the author provides their perspective on a research area, topic, or issue. They are intended to provide contributors with a forum through which to air their own views and experiences, with few of the constraints that govern standard research articles. This Food for Thought article is one in a series solicited from leading figures in the fisheries and aquatic sciences community. The objective is to offer lessons and insights from their careers in an accessible and pedagogical form from which the community, and particularly early career scientists, will benefit.
The International Council for the Exploration of the Sea (ICES) and Oxford University Press are pleased to make these Food for Thought articles immediately available as free access documents

on the team were not much younger than me, and they knew all the best places on the island to explore—especially in and around the ocean. Had I stayed a few years longer at Punahou, a young Barry Obama might have shown up in my class. However, I soon moved to the outer islands so I could devote more time to my new passions: body surfing and skin diving. By 1974, I was living in Australia, where I spent as much time as possible under the waters of the Great Barrier Reef and evolved from hunting fish with a speargun to stalking them with a camera. My first publication was actually an underwater photograph that appeared in *The Ocean World of Jacques Cousteau*.

A popular saying that emerged from Berkeley in the 1960s was "Don't trust anyone over 30" (Bartleby.com, 1989). When I reached that pivotal age in 1977, I was back in the United States following a brother's wedding and finding it difficult to ignore the WSIDWIGU question any longer. I considered going back to teaching, but demographics were unfavourable: the baby-boomer surge was over and teachers were being laid off. Plan B was to become a marine biologist, but that option was problematical because the only science course I had taken in college was Astronomy. Fortunately, I was still a Hawaiian resident, so I enrolled at the University of Hawaii and began 2 years of undergraduate science courses so I could be a credible applicant to graduate school.

And here is where serendipity first intervened: my randomly assigned curriculum advisor was Jim Shaklee, who taught ichthyology and happened to be a wizard at using protein electrophoresis to study allozymes (variant forms of enzymes, distinguished based on differing sequences of amino acids). I took Jim's excellent course and, short of money, asked about a summer job; he did not have a "job", but I could do a volunteer project in his laboratory and he would teach me electrophoresis. (The financial problem was solved with another stroke of luck; my chemistry laboratory partner was leaving town and I inherited his lucrative, night-time job in Waikiki serving wine to sunburned tourists.) The project we settled on, suggested by Jack Randall at the Bishop Museum, was to figure out whether two forms of a shallow-water lizardfish (genus *Synodus*) represented one or two species. This was a perennial problem in systematics: do observed differences between morphotypes merely reflect natural levels of variation within a species? Or are the forms fundamentally different types of organism? Protein electrophoresis has clear potential to inform this type of problem: amino acids are coded for by sequences of DNA bases (the "genetic code"), so different allozymes can be inferred to reflect different genes. The idea of using a new genetic method to solve a puzzle that had stumped world-famous ichthyologists (Jordan and Evermann 1905; Gosline and Brock 1960) was exhilarating to a neophyte scientist.

The lizardfish were relatively easy to collect, and soon I had several specimens of each morphotype. The procedure involved saturating wicks of filter paper with extracts from four tissues (muscle, liver, heart, eye), loading them onto gels made from potato starch, and subjecting them to electric current. Proteins with different amino-acid sequences migrated in different directions and/or at different rates, and at the end you could visualize where each sample had migrated to by bathing a slice of the gel in a solution that contained reactants for the enzyme and a linking dye. This created dark bands on the gel that magically appeared before your eyes as the reaction progressed. When bands on the first set of gels began to emerge, the result was unmistakable: two species!—the two morphotypes differed consistently at about half their genes! At that moment I had the answer to my WSIDWIGU question: I wanted to study the systematics and population genetics of fish. And that is largely what I have done for the last ~40 years.

But what would my scientific career be like if I had been assigned a different curriculum advisor at University of Hawaii? Even if I had remained interested in marine science, the focus of my research (and hence what I ended up working on) likely would have been quite different.

Uncontrolled controls

Chance interceded again soon after I started the lizardfish project. Jim Shaklee impressed upon me the importance of having a control on each gel that would produce bands of known mobility. For the controls I chose a common lizardfish in another genus, *Saurida gracilis*, which all authorities agreed was a single, polytypic species distributed widely throughout the Indo-Pacific. I collected *Saurida* from various habitats around Oahu, and each time I ran *Synodus* experiments I took a new *Saurida* specimen out of the freezer, extracted tissue samples, and loaded them on the same gels. This produced some puzzling results. When I did replicate samples of the same *Synodus* specimens on different days, the relative positions of the *Synodus* and control bands did not remain constant. After many frustrating days spent resampling numerous *Synodus* specimens failed to resolve the problem, for lack of a better idea I decided to examine the controls in more detail by loading multiple *Saurida* specimens on the same gels. This produced an astounding result: the *S. gracilis* "controls" represented not one, not two, but three distinct species, each characterized by fixed differences at ten or more genes! Once the various specimens could be sorted genetically into three groups, it was apparent that they had finely partitioned the near-shore marine environment: specimens in group A were all found in very shallow (<2 m) brackish or muddy water; those in group B occurred a little deeper (generally 3–10 m) on coral or nearby sandy patches; and specimens in group C were found in similar habitats to group B, but generally at depths of 10 m or more.

Given this startling finding, I set aside the *Synodus* project to focus on *Saurida*. With the three groups of specimens unambiguously defined genetically, I found that what appeared to be continuous variation at several morphometric characters turned out to be mixtures of discrete or semi-discrete distributions. For example, in the overall collection of ~70 *Saurida* individuals, the number of pelvic fin rays ranged from 11 to 15, with the type of bell-shaped distribution that often characterizes variation within a species (Figure 1a). It turns out, however, that most individuals from group A have 12 fin rays, most from group B have 13, and most from group C have 14 (Figure 1b). After defining the groups based on independent genetic characters, multivariate morphometric data could be used to classify new individuals (e.g. type specimens from museums) for which genetic data were not available. Using this approach, I was able to show in my first scientific paper (Waples, 1981) that group B represented the "true" *S. gracilis*, that specimens from group A belonged to a species (*Saurida nebulosa*), that long ago (mistakenly, it turns out) had been synonomized with *S. gracilis*, and that group C individuals were new to science and hence were given the name *Saurida flamma*, after the flame-coloured bands on their mouth. All of this only transpired as a chance consequence of deciding to use

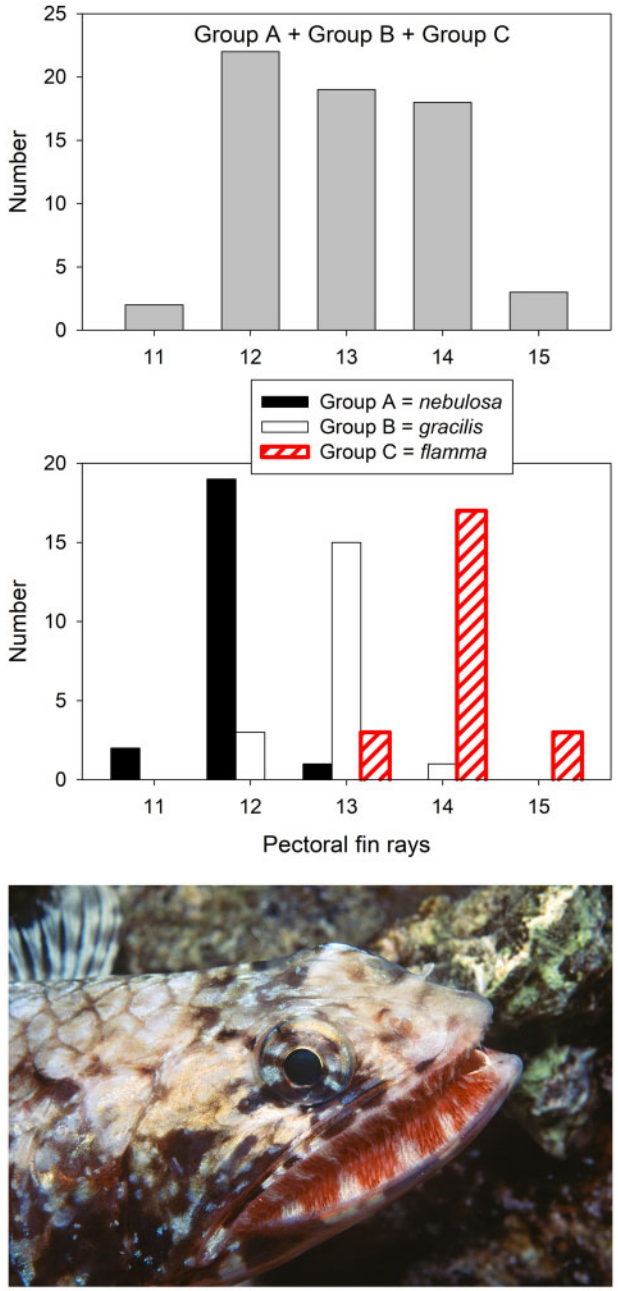

Figure 1. Top: distribution of pectoral fin ray counts for a collection of shallow-water lizardfishes (genus *Saurida*) from Hawaii, all reputed to be *Saurida gracilis*. Middle: after genetic analyses sorted the specimens into three groups (A, B, C), the overall distribution of pectoral fin rays was revealed to be a mixture of three semi-discrete distributions. Subsequent analyses revealed the valid scientific names for each group, as shown in the legend. Bottom: the new species, *Saurida flamma* (group C), named for the flame-coloured bands on its mouth. Photo copyright Keoki Stender, by permission.

this common species as a control in a study designed for other purposes.

Eventually I did return to the *Synodus* problem and teamed up with Jack Randall to revise the genus in Hawaii, including descriptions of four new species (Waples and Randall, 1988).

Confronting your data

Even under the best of circumstances biological data are messy, so luck as well as skill is usually involved when clear patterns emerge. Artefacts can creep into your data in so many ways that scepticism and careful scrutiny are needed at all steps of data collection and analysis. One small data problem I discovered by chance illustrates this pervasive issue. In 1986, I arrived in Seattle to work as a postdoc for the US National Marine Fisheries Service (NMFS; aka NOAA Fisheries) and soon started a genetic monitoring programme for Snake River Chinook salmon (*Oncorhynchus tshawtscha*) and steelhead (anadromous *Oncorhynchus mykiss*) that is still ongoing (Waples *et al.*, 1993; Van Doornik *et al.*, 2013). At that time, monitoring involved collecting samples for protein electrophoresis from young-of-the-year juveniles. For the smallest fish (<60 mm), it was difficult to get enough heart and eye to produce reliable results, often leading to missing data for some of the 30–40 variable gene loci. On allozyme gels, heterozygotes are often harder to score because their activity is spread across two or more bands and they can be easily missed with poor-quality samples.

For a project focusing on Chinook salmon, I wanted to see whether the amount of missing data in an individual was correlated with its heterozygosity, which would indicate a potential bias. I did not find a significant heterozygosity-missing data correlation, but I did find an odd result. Most individuals had no missing data, some were missing data for 1–3 loci, and a few had as many as five or six missing data points, with the latter group largely explained by the failure of one tissue to produce viable results. Curiously, I also found a small group of individuals with missing data for 21 or 22 gene loci. What could cause such a result? It was not due to complete failure of experiments to work on a given day, as these odd individuals were sprinkled among other normal individuals on many gels assayed at different times. I pulled some of the odd specimens from the freezer and, after a closer examination, it was clear that they were steelhead, not Chinook salmon. The two species are generally easy to distinguish, but small specimens can be mistaken for each other during field collections. But steelhead and Chinook salmon have been separated for ~10 million years and have substantial genetic differences, so how could this contamination not have been noticed before? A check with the laboratory staff—all of whom had several years of experience scoring salmon gels—clarified that the offending specimens had been noted, but the way the data were recorded did not reflect that. When species X was being analysed and one or more oddballs appeared on a gel, the laboratory staff would note that "aha, that is an allele from species Y". At that point, the oddballs would be "zeroed out", meaning that their data were recorded as missing—but only for the gene loci currently being assayed. There was no formal procedure to flag those individuals as imposters and systematically remove all their data. This meant that, for the substantial fraction of loci for which species X and species Y shared alleles, individuals of species Y would be scored as if they were species X. The consequence of this was that frequencies of common alleles shared by the two species would be inflated for species X, whereas frequencies of rare alleles at the same loci would be underestimated.

As a result of this chance discovery, a more robust procedure was implemented to deal with inadvertent mixtures of species in field samples. In this case, the practical consequences of the errors were limited, as only a small fraction of the samples were steelhead rather than Chinook salmon and the effect on estimated

allele frequencies was modest. However, this represents only one small example of more general issues about data quality that can have profound consequences. For example, a paper published in *Proceedings of the National Academy of Sciences of the United States of America* (Ottmann *et al.*, 2016) claiming that sibling groups of larval rockfish (*Sebastes* sp.) travel together for many months was later retracted (see https://www.pnas.org/content/114/52/E11336) when it was discovered that the small groups of "siblings" were actually a different species of rockfish. Compared to differences between species, differences among the few specimens of the second species were so small that the analyses the authors performed concluded that they had to be siblings. They had not checked for the presence of multiple species in their samples, even though 60 species of *Sebastes* are found along the west coast of the United States.

Irrepressible effects of N_e

It is often said that effective population size (N_e) is one of the most important parameters in evolutionary biology, and I have been known to make similar statements myself. That might be true, but N_e is also insidious because it manifests itself in diverse and unexpected ways—as was made clear to me soon after I moved to Seattle. A major focus of our laboratory was the analysis of mixed-stock salmon fisheries using a form of genetic mixture analysis (Utter *et al.*, 1987; Shaklee *et al.*, 1999). I had become interested in a related problem, which was whether linkage disequilibrium (LD; non-random associations of alleles at different gene loci) could be used to detect the mixtures of salmon populations. It was known that samples that include individuals from more than one population generate mixture LD, with magnitude that depends on the mixture fraction and how large the genetic differences are among populations (Nei and Li, 1973). Following interbreeding, the LD signal decays over time but still is potentially detectable for several generations.

I wanted to know what statistical power one might have to detect salmon mixtures using LD, given empirical data on genetic differences among Chinook salmon populations. For this effort, I recruited the help of Peter Smouse, who had done seminal work on mixture LD in indigenous tribes from South America (Smouse and Neel, 1977). We simulated many *in silico* mixtures of salmon populations, and because we explicitly modelled reproduction, we had to stipulate a population size. We used a Wright–Fisher random reproductive success model, modified to account for salmon age structure, and with a wide range (20-fold) of effective population sizes. As expected, we found reduced statistical power for weakly differentiated populations, longer time after the interbreeding event, and unequal mixture fractions. Unexpectedly, however, we also found that these patterns were often dwarfed by the N_e effect, with smaller effective sizes producing more LD and much higher statistical power (Waples and Smouse, 1990; Figure 2). By three generations of random mating following an admixture event, most of the remaining LD in the mixed population could be attributed to drift.

This unexpected result alerted me to the powerful effects of N_e on population genetic data, a theme I have pursued in many subsequent studies. With a little detective work, I found a paper by Bill Hill (1981) that showed how one could estimate N_e based on the amount of LD in a sample and a paper that described a similar approach based on temporal changes in allele frequency (Nei and Tajima, 1981). At the time, it was thought that genetic methods for

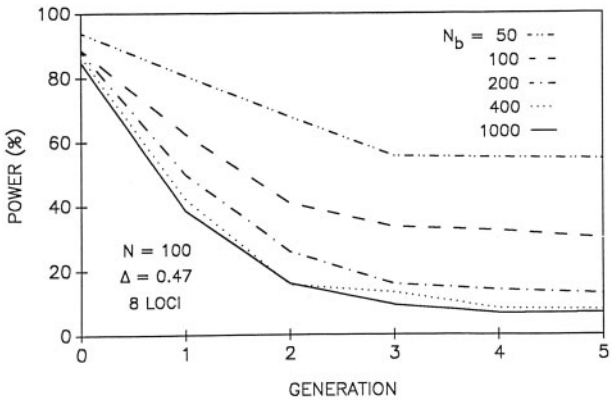

Figure 2. Power to detect mixtures of salmon populations based on linkage disequilibrium (LD). across all pairwise combinations of eight gene loci. Power is the percentage of tests of the null hypothesis (H_0: LD = 0) that are rejected. Results are from computer simulations that randomly interbred equal numbers of individuals from two populations in generation 0. Δ is a measure of genetic differentiation between the populations, and sample size was $N = 100$ individuals. Power is shown as a function of the number of generations following the interbreeding event and the effective number of breeders each year in each simulated population (N_b). Reproduced from Waples and Smouse (1990).

estimating N_e would only be useful for species (e.g. *Drosophila*) with populations too large to enumerate. However, the genetic signal these methods are sensitive to is proportional to $1/N_e$, which means that the methods have more precision for small populations. I had ideas about how to refine these methods for application to species of conservation concern and was just starting to develop them (Waples, 1989; Waples and Teel, 1990). At this point, however, chance intervened once again to change my career trajectory, to such an extent that any further efforts to pursue the study of genetic estimators of N_e had to be deferred for over a decade.

Endangered Species Act interludes

In 1990, the lid was about to blow off the pressure cooker of salmon conservation and management in the US Pacific Northwest. A decade before, concerns about declining salmon populations were forestalled for a time by passage of the Northwest Power Act, which led to development of the Columbia River Basin Fish and Wildlife Program, at the time considered "the most ambitious and costly effort at biological restoration on the planet" (Lee and Lawrence, 1985, p 433). A major goal of the Fish and Wildlife Program was to double salmon and steelhead abundance within 10 years. Unimpressed by this prediction, the salmon populations themselves continued to decline, and in 1990, early drafts of a report (subsequently published as Nehlsen *et al.*, 1991) documenting over 200 at-risk salmon stocks were circulating within the region.

That year our agency, which has stewardship responsibility for marine and anadromous species under the US Endangered Species Act (ESA), received petitions to list several groups of Columbia River Basin salmon populations as threatened or endangered under the ESA. To that point I had paid little attention to the ESA and did not even realize that the Act affords legal protection to any entity that meets the ESA's definition of "species", which includes named subspecies and (for vertebrates) "distinct

population segments" (DPSs). Grizzly bears, bald eagles, and alligators were listed as DPSs in the contiguous United States, even though they were more abundant elsewhere, and the ESA salmon petitions sought listings based on the DPS provision. But the term "DPS" does not have a clear biological meaning, and the ESA provides no guidance on how to identify DPSs. When NMFS policy staff asked our laboratory for scientific guidance on the validity of the ESA salmon petitions, my response was that someone needed to define what a DPS of salmon was. One thing I knew for sure: it was not my responsibility, as my job description said nothing about the ESA. The problem was, in 1990, nobody at our Center had anything related to the ESA in their job description, yet somebody had to take charge of the issue. How that someone became me hinged on another random event.

The ability to list population-level units was provided for in the original (1973) implementation of the ESA, and the current DPS language dates from 1978 amendments. In 1990, most of the DPS listings (including those for the three iconic species mentioned above) had been carried out by the US Fish and Wildlife Service (USFWS), which has ESA responsibility for terrestrial species, but each DPS listing determination had been done on an *ad hoc* basis, with no formal policy guidance, so the record of past determinations did not allow one to predict with any certainty how the agencies might handle a new DPS evaluation in the future. This was particularly troublesome for salmon: each of the Pacific salmon species comprises many hundreds of separate populations, which potentially could be grouped into DPSs in a nearly infinite number of different ways.

To address the lack of consistency in prior DPS determinations, and to deal with the Pandora's Box of endangered species issues for Pacific salmon that had just been opened up, in June 1990, the USFWS and NMFS convened a workshop in Washington, DC. The goals were to develop (i) an overarching DPS policy that would apply to all species and (ii) consistent with provisions of the broader policy, more specialized guidance for the complex issues involving salmon. Attendees included population biologists and geneticists, policy staff, and lawyers from both agencies. Another geneticist from our center was invited to attend but had a conflict, so I was sent instead. Following the meeting, efforts by USFWS to develop a broad DPS policy faltered. However, the 1990 ESA salmon petitions had tight legal deadlines and our agency could not afford to wait to develop a framework for evaluating salmon DPSs. Because I had attended the Washington, DC, meeting and was working on salmon, I was fingered as a likely suspect to draft a scientific paper that would, it was hoped, form the basis for formal policy guidance by our agency.

Lawyers at the DPS workshop provided some useful background information for context: (i) legislative discussions leading up to passage of the ESA made it clear a major goal was conservation of biodiversity; (ii) the ESA itself stipulates that listing determinations be scientifically based; and (iii) it was recognized that the ability to list populations could be abused (if, for example the squirrels in a city park were listed as a DPS because they were isolated from other squirrels by urbanization), so the agencies were directed to use the DPS provision sparingly. Early drafts emerging from USFWS following the workshop were unfocused and provided a laundry list of options, reflecting the range of views discussed at the meeting: a DPS might be a, or b, or c, or d, or and so on.... I did not see how this would provide any meaningful guidance to future users, nor did I see how it would resolve the "squirrels in a city park" issue. It seemed to me that what was needed was a simple, two-part test with the criteria joined by "and" rather than "or". It was clear that a "distinct" population segment must involve substantial reproductive isolation, but by itself that is not sufficient, as city park squirrels might meet that test. The need for a second criterion becomes obvious if one thinks about one of the major goals of endangered species conservation—avoiding extinctions because they are irreversible. Extinctions are permanent because they represent loss of the genetic blueprint for making a specific type of organism. Therefore, to satisfy the second criterion, a DPS should represent a major component of genetic diversity within the species as a whole.

Accordingly, I drafted a scientific document outlining these ideas, fleshed out with several practical considerations for application to salmon. I had no idea how these things worked within government agencies; I expected the draft to disappear somewhere in the beaurocracy and never be seen again. But the draft was received well locally and also at headquarters—and even the lawyers seemed to like it. After the paper was favourably peer reviewed, a draft ESA salmon policy based on the science paper was published (NMFS, 1991a) and this was used to guide responses to the 1990 salmon petitions. After public review and comment, both the science paper (Waples, 1991) and the salmon DPS policy (NMFS, 1991b) were finalized. Several years later, a joint interagency DPS policy was finally published (USFWS and NMFS, 1996), which employed a two-part test similar to that used to define salmon DPSs.

This was just the start of my salmon-ESA involvement. I became head of a group of scientists charged with the decade-long task of developing the scientific basis for ESA listing determinations for all US West Coast species of Pacific salmon, as well as steelhead and anadromous cutthroat trout, *Oncorhynchus clarkia*. At the same time, we had to evaluate the likely consequences of various management/conservation actions, such as operation of dams and hatcheries, that could affect listed populations. In the early-2000s, we initiated formal ESA recovery planning for listed populations and formed a series of teams to develop science-based recovery goals consistent with long-term viability. These were all-consuming tasks that occupied ~150% of my time. Eventually, although I approached the Event Horizon more than once, I was able to escape the powerful gravitational pull of the ESA—at least to the extent that I could focus on other things for a change.

Parentage analysis without parents

The most pleasant serendipitous outcome in my scientific career has been the opportunity to collaborate and publish with my son, Ryan, but the path that led to that result was far from linear. I did not try to steer my offspring's interest towards (or away from) science. However, one summer, faced with the prospect of a bored teenager moping around the house, I brought home my copy of *A Primer of Population Biology* (Wilson and Bossert, 1971) and suggested that Ryan might want to look it over. He gave a non-committal grunt but later acknowledged that he did read the book and that it might even have played a small role in shifting his main interest from chemistry to biology, which he ended up majoring in college. Even during his undergraduate years, but especially after he graduated and spent several years claiming to be considering applying to graduate school, I tried to get Ryan to learn computer programming as an essential skill of a

modern biologist. Like many parental suggestions, this one was routinely ignored until I repackaged it as an interesting problem and challenged him to find the answer through simulations.

The problem arose from a family holiday tradition. On an agreed-upon day, we all met after breakfast to pick names out of a hat. We then dispersed, each to buy a gift for the randomly chosen family member, and reconvened at lunch to exchange gifts. Sometimes one of us drew our own name, so we had to redo the draw. The problem I posed to Ryan: figure out the probability that at least one person will draw their own name, and how that probability changes with the number of names in the hat, n (which varied depending on whether Ryan or his sister Jade had a significant other at the time). The probability of having to redraw is easy to work out by hand for small numbers (1/2 for $n = 2$; 2/3 for $n = 3$; 5/8 for $n = 4$), but this rapidly becomes very tedious for $n > 5$.

At the time I programmed in several Stone-Age languages (Fortran, Pascal, Basic), but Ryan decided to teach himself Python, which was a fortuitous choice that facilitated his subsequent forays into bioinformatics. Before long, he had produced results: after some gyrations for small numbers of participants, the probability of having to redraw converges rapidly on a value a bit over 0.63 (Figure 3). That value seemed curious, but when I mentioned this exercise to my mathematician brother-in-law, he was excited: "This is the famous hat-check problem! All the gentlemen going to the theater check their hats and get a ticket with a number, but after the performance the attendant is nowhere to be found, so hats are passed out at random. The probability that at least one person gets their own hat converges on $1 - 1/e \approx 0.6321$."

Before long, Ryan's new programming skills came in handy in relation to a topic that had attracted my attention, which was understanding the evolutionary responses by salmon to major anthropogenic changes to their environments (Waples *et al.*, 2008, 2017). In the late-2000s, colleagues and I collaborated on a salmon and climate change project, a core feature of which was developing an individual-based model that allows for both

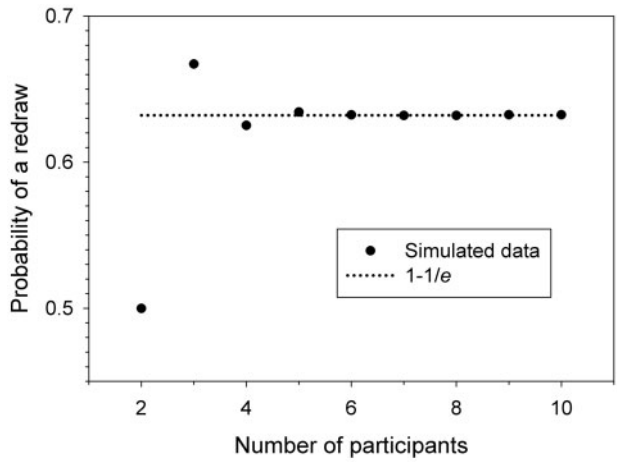

Figure 3. The gift-exchange (aka "hat-check") problem. A hat contains the names of n participants, who must buy a gift for the person with the name they draw from the hat. A redraw is required if anyone draws their own name. What is the probability of a redraw, and how does it change with n? Filled circles are from numerical simulations; the dotted line shows the theoretical expectation that the probability converges on $1 - 1/e$ for large n.

evolution and phenotypic plasticity (Reed *et al.*, 2010, 2011). This model included an option to have the amount of additive genetic variance in a population decline if N_e dropped <500—in accordance with the "50–500" rule, the later part of which holds that an effective size of ~500 is needed to ensure that long-term loss of genetic variation by drift is balanced by the creation of new variation by mutation (Franklin, 1980).

I wanted to test how this option was working in the model by tracking both N_e and additive genetic variance over time. The standard textbook formula for computing inbreeding effective size (Crow and Denniston, 1988) depends on three parameters: the number of potential parents (N) and the mean ($\bar{k}$) and variance (V_k) in number of offspring per parent:

$$N_e = \frac{\bar{k}N - 2}{\bar{k} - 1 + {}^{V_k}/_{\bar{k}}} \quad (1)$$

The way our programme was coded, however, it was not easy to identify all potential parents in a given generation and count the number of offspring produced by each that survived to return as adults. However, it was easy to "ask" each offspring who its two parents were, and by integrating across all offspring one could calculate $\bar{k}$ and V_k. But this method only provided information about parents that actually produced at least one offspring. What about the parents that produced no offspring? These null parents should be included in total N, but how many of these were there? The method I used to calculate the mean and variance of offspring number provided no information about this class of parents.

To assess the effects of null parents on inbreeding N_e, I created offspring distributions in which some parents, by chance, produced zero offspring. I then calculated N_e using (1), both with and without the null parents. Results were surprising: there was no effect of null parents. Whether null parents were included or not affected all the key parameters (N, $-k$, and V_k), but they changed in such a way that N_e was unchanged. To explore this analytically, I took simple expressions for the mean and variance ($\bar{k} = \Sigma k_i/S$ and $V_k = \Sigma k_i^2/S - \bar{k}^2$) and inserted them into (1). After a little rearrangement, the formula for inbreeding N_e reduces to this simple expression (Waples and Waples, 2011):

$$N_e = \frac{2S - 2}{\frac{\sum k_i^2}{2S} - 1}, \quad (2)$$

where k_i is the number of offspring produced by the i^{th} parent and S is the number of offspring that have been assigned to parents. Expressed this way, it is clear that inbreeding N_e does not depend on N; the only unknown in (2) is Σk_i^2 = the sum of squares of offspring number. Null parents make no contribution to Σk_i^2 and hence none to inbreeding N_e. This simple relationship does not hold for variance N_e, except in the special case where $S = N$.

Although genetic methods to assign offspring to parents are now routinely applied to natural populations (Jones *et al.*, 2010), analyses become complicated when only some of the potential parents can be sampled. Equation (2) is quite versatile in this respect, as it is not affected by either the number of parents or the number sampled. I enlisted Ryan's help to develop an algorithm to infer the vector of parental contributions (the k_i values), given a set of correctly specified sibling

relationships. This allowed us to calculate Σk_i^2 and hence N_e from (2) without sampling any parents at all—that is, by conducting parentage analysis without parents. Ryan was also able to show that, for a given set of inferred sibling relationships, the estimate of N_e obtained using (2) is the same as that produced by Wang's (2009) sibship method (Ackerman *et al.*, 2017).

Equation (1) or similar versions had been in widespread use for over a half century, but I only stumbled on the simpler and quite useful (2) when it was not possible to obtain the data I wanted by conventional means. This exercise was enriched by the programming skills of Ryan, which trace their origins to curiosity about a problem related to holiday gift exchanges.

Conclusion

Sometimes Lady Luck grabs you by the throat and can direct your life for a decade or more; for me, the chance assignment of Jim Shaklee as my advisor and the fateful 1990 meeting about the ESA that I was not supposed to attend fall into that category. When opportunities like this arise, you should be prepared to step up and make the most of them. In baseball parlance, when Fortune hangs a curveball in your wheelhouse, jump on it! But sprinkled throughout a life are many more chance events that make only small ripples and are easy to miss if you are not paying attention. Several occurrences like this that have enriched my scientific career are described above. Although these events are difficult to categorize and hard to generalize about, in many cases they present initially as annoying problems (e.g. the poorly behaved controls on the *Synodus* gels; the cluster of individuals missing data for about half their gene loci; the powerful effects of genetic drift that complicated the intended analyses; the difficulty in getting computer code written by someone else to produce the specific output you want). Often there is a silver lining to these annoyances, if one only takes the time to look. In situations like this, one can hope to increase the chances of a serendipitous outcome by adopting the philosophy of Niels Bohr: "How wonderful that we have met with a paradox. Now we have some hope of making progress". It is not reasonable to expect that you will be rewarded with insights as momentous as those of Bohr regarding Quantum Mechanics, but consistently adopting this perspective can lead to important contributions in the long run. This is particularly true in the analysis of empirical (or even simulated) data. All the time you can muster to poke, prod, and examine your data from every possible angle will often be repaid by the discovery of anomalies that would not have been noticed otherwise. At worst, you might identify a problem and improve the quality of your data. If you are lucky, you will discover something new that leads down a novel and interesting path. Although I could retire at any time, I am still active in research because I am still discovering new things and still learning from collaborations with (mostly) younger scientists—and who knows when serendipity might manifest itself again?

References

Ackerman, M. W., Hand, B. K., Waples, R. K., Luikart, G., Waples, R. S., Steele, C., Garner, B. A. *et al.* 2017. Effective number of breeders from sibship reconstruction: empirical evaluations using hatchery steelhead. Evolutionary Applications, 10: 146–160.

Barlteby.com. 1989. Jack Weinberg. Respectfully Quoted: A Dictionary of Quotations. https://www.bartleby.com/73/1828.html (last accessed 24 February 2019).

Crow, J. F., and Denniston, C. 1988. Inbreeding and variance effective population numbers. Evolution, 42: 482–495.

Franklin, I. R. 1980. Evolutionary change in small populations. In Conservation Biology: An Evolutionary–Ecological Perspective, pp.135–150. Ed. by M. E. Soule´ and B. A. Wilcox Sinauer Associates, Sunderland, MA, USA.

Gosline, W. A., and Brock, V. E. 1960. Handbook of Hawaiian Fishes. Univ. Hawaii Press, Honolulu.

Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genetical Research, 38: 209–216.

Jones, A. G., Small, C. M., Paczolt, K. A., and Ratterman, N. L. 2010. A practical guide to methods of parentage analysis. Molecular Ecology Resources, 10: 6–30.

Jordan, D. S., and Evermann, B. W. 1905. The aquatic resources of the Hawaiian Islands. Part I. The shore fishes. Bulletin of the United States Fish Commission, 23: 1–575.

Lee, K. N., and Lawrence, J. 1985. Adaptive management: learning from the Columbia River Basin Fish and Wildlife Program. Environmental Law, 16: 431.

Nehlsen, W., Williams, J. E., and Lichatowich, J. A. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. Fisheries, 16: 4–21.

Nei, M. and Tajima, F. 1981. Genetic drift and estimation of effective population size. Genetics, 98: 625–640.

Nei, M., and Li, W.-H. 1973. Linkage disequilibrium in subdivided populations. Genetics, 75: 213–219.

NMFS. 1991a. Federal Register, 56: [13 March 1991]:10542.

NMFS. 1991b. Federal Register, 56: [20 November 1991]:58612.

Ottmann, D., Grorud-Colvert, K., Sard, N. M., Huntington, B. E., Banks, M. A., and Sponaugle, S. 2016. Long-term aggregation of larval fish siblings during dispersal along an open coast. Proceedings of the National Academy of Sciences of the United States of America, 113: 14067–14072.

Reed, T. E., Waples, R. S., Schindler, D. E., Hard, J. J., and Kinnison, M. T. 2010. Phenotypic plasticity and population viability: the importance of environmental predictability. Proceedings of the Royal Society of London, Series B, 277: 3391–3400.

Reed, T. E., Schindler, D. E., Hague, M. J., Patterson, D. A., Meir, E., Waples, R. S., and Hinch, S. G. 2011. Time to evolve? Potential evolutionary responses of Fraser River sockeye salmon to climate change and effects on persistence. PLoS One, 6: e20380.

Shaklee, J. B., Beacham, T. D., Seeb, L., and White, B. A. 1999. Managing fisheries using genetic data: case studies from four species of Pacific salmon. Fisheries Research, 43: 45–78.

Smouse, P. E., and Neel, J. V. 1977. Multivariate analysis of gametic disequilibrium in the Yanomama. Genetics, 85: 733–752.

USFWS and NMFS. 1996. Federal Register 61 [7 February 1996]: 4722.

Utter, F., Teel, D., Milner, G., and McIsaac, D. 1987. Genetic estimates of stock compositions of 1983 chinook salmon, *Oncorhynchus tshawytscha*, harvests off the Washington coast and the Columbia River. Fishery Bulletin, 85: 13–23.

Van Doornik, D. M., Eddy, D. L., Waples, R. S., Boe, S. J., Hoffnagle, T., Berntson, E. A., and Moran, P. 2013. Genetic monitoring of threatened Chinook salmon populations: estimating introgression of non-native hatchery stocks and temporal genetic changes. North American Journal of Fisheries Management, 33: 693–706.

Wang, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. Molecular Ecology, 18: 2148–2164.

Waples, R. S. 1981. A biochemical and morphological review of the lizardfish genus Saurida in Hawaii, with the description of a new species. Pacific Science, 35: 217–235.

Waples, R. S. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics, 121: 379–391.

Waples, R. S. 1991. Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. Marine Fisheries Review, 53: 11–22.

Waples, R. S., Elz, A., Arnsberg, B. D., Faulkner, J. R., Hard, J. J., Timmins-Schiffman, E., and Park, L. K. 2017. Human-mediated evolution in a threatened species? Juvenile life-history changes in Snake River salmon. Evolutionary Applications, 10: 667–681.

Waples, R. S. O. W., Johnson, P. B., Aebersold, C. K., Shiflett, D. M., VanDoornik, D. J., Teel, A. E. and Cook, 1993. A genetic monitoring and evaluation program for supplemented populations of salmon and steelhead in the Snake River Basin. Annual Report of Research, Bonneville Power Administration, Portland. 179 pp.

Waples, R. S., and Randall, J. E. 1988. A revision of the Hawaiian lizardfishes of the genus *Synodus*, with descriptions of four new species. Pacific Science, 42: 178–213.

Waples, R. S., and Smouse, P. E. 1990. Gametic disequilibrium analysis as a means of identifying mixtures of salmon populations. American Fisheries Society Symposium, 7: 439–458.

Waples, R. S., and Teel, D. J. 1990. Conservation genetics of Pacific salmon. I. Temporal changes in allele frequency. Conservation Biology, 4: 144–156.

Waples, R. S., and Waples, R. K. 2011. Inbreeding effective population size and parentage analysis without parents. Molecular Ecology Resources, 11: 162–171.

Waples, R. S., Zabel, R. W., Scheuerell, M. D., and Sanderson, B. L. 2008. Evolutionary responses by native species to major anthropogenic changes to their ecosystems: pacific salmon in the Columbia River hydropower system. Molecular Ecology, 17: 84–96.

Wilson, E. O., and Bossert, W. H. 1971. A Primer of Population Biology. Sinauer Associates, Sunderland, MA, USA.

Handling editor: Howard Browman