What Does Genetics Have to Do with It?

- AA
- aA
- Aa
- aa
1

Even Fish Obey Mendel’s Laws

A.J. Gharrett
University of Alaska Fairbanks, School of Fisheries and Ocean Sciences, Fisheries Division, Juneau, Alaska

This is the first of a series of articles that consider the role genetics plays in conservation and management of our fisheries resources. Genetics is the discipline that makes a coherent field out of many diverse biological disciplines, which include the molecular level (e.g., biochemistry and cell biology), the organism (e.g., anatomy and physiology), interactions among organisms (e.g., ecology and population biology), and descriptions and phylogenetic relationships of species (e.g., zoology and ichthyology). Inheritance, its underlying mechanisms, and ultimately evolution, which at its simplest level is merely changes in the genetic instructions, are the basis of all biological disciplines. Unfortunately, the key to the door to genetics has a price. A number of underlying ideas are important to understanding the role of genetics. In addition, a special vocabulary (jargon if you are not already a geneticist) is used to describe these ideas. The first two chapters deal with the basic principles and vocabulary, but we will add more terms along with ideas introduced in later chapters. In this first chapter, we review basic inheritance, which is often referred to as Mendelian genetics.
“What does genetics have to do with it?” That’s one of the most frequent questions that fisheries geneticists get asked by fish culturists, managers, fishermen, and the public. Some of the more specific questions are, “Do hatcheries change the fish? If they do, is that always bad?”; “Do hatchery fish harm wild stocks?”; “Does regulating fisheries by time, place, gear, or whatever affect the stock in the long term?” It turns out that the universal answer is, “It depends” because nearly every situation is different, and as the adage goes, the devil is in the details. Answers to these apparently simple questions require a thorough understanding of the biology involved in the situation, which includes the pertinent genetics issues. Often the total information that is needed for an informed answer or decision exceeds what is available to culturists, managers, and geneticists combined. Consequently, fisheries geneticists, like fisheries scientists, often have to make the best decision possible from incomplete or inadequate data. In such instances, it is usually preferable to err on the side of caution; that is, make a decision that is least likely to harm the resource. This is sometimes referred to as the precautionary principle and, as for physicians, the first rule is “do no harm.” Just like the decisions physicians make, those of managers and geneticists may be irreversible.

In order to consider these and other questions, we need to learn a little about basic genetics. Our ultimate goal is to examine the underlying genetic principles and then to apply the principles to common questions about the role or influence of genetics on fisheries management and conservation. One of the barriers to understanding genetics is the jargon that geneticists use. Along the way we will attempt to clarify some of that jargon because those explanations will also be useful in revealing many of the basic concepts. We will refer to this chapter (1. Even Fish Obey Mendel’s Laws) in other chapters.

Local adaptation

One of the central concepts that provides us with a perspective in understanding how genetics applies to conservation and management of wild populations, and the interaction of wild and cultured stocks, is local adaptation. If one looks at the spawning and rearing habitat (for example) of different salmon populations throughout a species range, it is apparent that no two streams are identical. For example, there are obvious latitudinal temperature changes—streams in the Pacific Northwest are generally warmer than streams in Alaska. And, even within larger drainages there may be variation in the available habitat—tributaries or reaches within tributaries may differ in gradient, stream flow, water chemistry, temperature regime, spawning substrate (gravel size and quality), and juvenile rearing conditions. In addition, populations of salmon species usually have different marine experiences. A clear example of local adaptation is the sockeye salmon in Necker Creek, which drains Benzeman Lake on southern Baranof Island, Alaska. Many fallen trees in Necker Creek block the passage of larger fish to the lake. Over the years, the size of spawners has been reduced by natural selection. The small fish that now spawn in the system are called “twofer’s” by the fishermen because it takes two Necker Creek fish to get the same price as a single sockeye salmon from other areas. Local adaptation has been documented in many other instances. One particularly striking result is spawning timing of sockeye salmon in the Fraser River system in British Columbia. In salmon, which are cold-blooded, development rate and time from spawning to hatching proceeds at a rate that depends on the temperature of their environment (i.e., water temperature). Embryos that are produced in warmer water generally develop faster and hatch earlier than those produced in colder water. There is a strong correlation between spawning timing and water temperature—Fraser River sockeye that spawn in coldwater drainages return earlier than those that spawn in warmer drainages.

Habitat diversity is also often apparent for populations of a marine species in different areas, although the differences may not be as obvious as the differences we identify in freshwater populations. Also, when environmental conditions are examined over time, it is clear that no two populations have experienced exactly the same sequence of historical environments. Moreover, because freshwater and marine conditions change from year to year, the lifetime experiences of different cohorts (year classes) produced by a single population often differ. And,
this does not even consider the changes in climate that inexorably occur over time. We will come back to the idea of local adaptation repeatedly, but for now let’s examine some basic genetics ideas and the jargon that is used so that we will be prepared to incorporate genetics into those discussions.

**Mitosis and cell division**

An analogy for the information encoded in the DNA of an organism is a set of blueprints or a recipe that details how to construct or cook a particular product. Blueprints detail how to build a house or a widget; recipes give instructions for baking casseroles or cakes. DNA specifies how to build a hemoglobin molecule or other biological structure, but also determines when and where in the organism the construction is to take place. In most animals, the vast majority of the tens of thousands of genes are carried on chromosomes that are located in the nucleus of a cell. We will also see later that we have many DNA sequences that do not carry information for structures. In fact, many sequences have no known function!

There is built-in redundancy for most genes because animals generally carry two separate copies of the blueprints. The redundancy serves several purposes. One is that there is a backup if one of the copies is defective. In some instances, both copies contribute to expression of a trait. Most of the organisms that we consider carry two copies of most genes. Organisms that usually carry two copies of genes are referred to as *diploid* organisms. Another benefit of carrying two copies is that the mechanism of sexual reproduction, which is based on the possession of two copies of each recipe, can generate an enormous number of different genetic combinations (design plans) within a species. The laws of genetics that Gregor Mendel discovered nearly 150 years ago apply to diploid genes that are carried in the nucleus. In organisms other than bacteria and blue-green algae, genes are arranged on chromosomes. Chromosomes are physical structures that are the vehicles which ensure that each of the two cellular products of a cell division (daughter cells) receives two complete (diploid) sets of genes (Figure 1). The chromosomes also make sure that *gametes* (eggs and sperm) receive exactly one copy of each gene. The single set or complement of chromosomes is referred to as a *haploid* set. Like the genes they carry, chromosomes occur in pairs. The idea is analogous to the pairs of animals on Noah’s ark. Each member of a pair of chromosomes carries information (genes) for the same traits and the genes are arranged in the same order on both chromosomes. A pair of chromosomes that carries information for the same traits is referred to as a *homologous* pair.

Cell division is necessary for fertilized eggs to develop into organisms, for organisms to grow, and for maintenance and repair. The process of cell division, which is called *mitosis*, ensures that complete diploid complements of chromosomes (and their genes) are received by each product of the division. Figure 1 shows how such exact partitioning is accomplished. The objective of the process is to distribute the chromosomes (and their genes) equally. Mitosis yields products that have identical diploid complements.

**Genotype versus phenotype**

An organism’s *genotype* refers to the genetic information that it carries, and the result, which we actually see in an individual, is the *phenotype*. For example, a recipe carries the information for baking a cake and is analogous to a genotype. The cake that results, however, depends on the quality of ingredients used, the skill of the baker, and occasionally the barometric pressure or altitude. No two cakes are identical, even though they were products of the same recipe. Some genes very explicitly define the resulting phenotype—like type AB blood—but others are not quite as explicit. For example, male pattern baldness in humans affects many people, but the extent of the effect runs from a healthy fringe to bald as a cue ball. The term genotype can be applied narrowly to the precise set of genetic information (for one gene) that exists for a particular trait, or somewhat more broadly for a specific combination of traits (such as temperature tolerance), or it can be used loosely to refer to the overall genetic composition. In the first case we might even specify the genotype by using notation. For example, let’s look at a genotype that includes three traits designated *A*, *B*, and *C*. They
**Mitosis (cell division)**

**Interphase** (resting cell) \((2N = 4)\)
1) DNA (chromosomes) extended: "bowl of spaghetti"
2) nuclear membrane intact
3) DNA replicates just prior to the initiation of cell division

**Prophase**
1) nuclear membrane breaks down
2) chromosomes thicken into short rods
3) centrosome breaks into centrioles
4) centrioles organize spindle apparatus

**Metaphase**
1) chromosomes line up at center
2) spindle fibers contract
3) centromere divides daughter chromosomes

**Anaphase**
1) contracting spindle fibers draw chromosomes to opposite poles

**Telophase**
1) cell pinches into 2 cells
2) spindle apparatus dismantles
3) chromosomes disaggregate
4) nuclear membrane reforms
5) centrioles become centrosomes

**Interphase**

Figure 1. Mitosis is the process that ensures that the products of cell division carry the same sets of chromosomes and their genes as each other and the parental cell.
could code for three separate traits like blood group and hair color or contribute to a single complicated trait like size. As an example, a particular organism might have two identical copies of gene $A$ for the first trait, two identical genes designated $b$ for the second trait, and two different versions of a gene, $C$ and $c$, for the third trait. The complete genotype would be $AAbbCc$. The $AA$ and $bb$ genotypes are referred to as homozygous (homo means the same) and the $Cc$ genotype is referred to as heterozygous (hetero means different).

Now we have a terminology problem: we have been using “gene” in two different ways. First, we used gene generally to describe the information that produces a particular trait, such as hair color. Second, we used it to refer to the different versions of the information for a particular trait, like red- or dark-colored hair. We can reconcile this problem by defining two more terms, locus and allele.

We mentioned previously that genes are carried by chromosomes. In fact, a gene for a particular trait is always found at the same site, or locus (= place), on a chromosome. A locus can refer to the location of a DNA sequence that codes for a protein like the components of hemoglobin, a DNA sequence that has no known function, or even a single nucleotide. My student, Mike, suggested a simple analogy to distinguish between loci and alleles: articles of clothing on an individual and their colors. In his analogy, locus is analogous to socks or shoes or gloves; alleles are the colors or types of the article. For example, Wanda has two blue socks, whereas Artie has a green sock on one foot and a red one on the other. Wanda would be homozygous at the sock locus and Artie would be heterozygous—possibly Artie is colorblind, but these hypothetical loci generally do not require any sense of style (Figure 2).

Multiple traits are referred to as loci, which is the plural of locus. The two (or more) slightly different versions of the DNA sequence at a locus are referred to as alleles. Again, the alleles might be alternative instructions for a particular trait, or just slightly different DNA sequences. In our clothing example, Wanda might have two identical black boots and Artie two different colored tennis shoes. To make a stronger connection with the genetics world (i.e., let’s get real), we have to understand that neither Wanda nor Artie may have had any choice about their garb because the clothes were provided by an oblivious governmental committee (indeed, they are probably lucky to have clothes). Artie and Wanda (and probably others) have different phenotypes. For some traits, the appearance or phenotype of a heterozygous individual ($Aa$) may be the same as the phenotype of the homozygous individuals ($AA$), but different from the others ($aa$).

For example if most people wore socks that had dark, unnoticeable colors and Artie had either one or two bright pink socks, he would soon be known as the guy with bright colored socks, whether he had one or two of them (Figure 3). A real example is albinism, absence of pigmentation, which occurs in many organisms. In most instances traits like albinism are referred to as a recessive trait and normal pigmentation is the dominant trait. So information for albinism is encoded at a particular locus (physical location) on a chromosome, and an individual
might have two alleles—slightly different versions of the instructions, one for normal pigmentation and another for albinism. An individual’s genotype could be either AA or Aa at the albinism locus, and its phenotype would be normal pigmentation. Only aa genotypes would be albino. The terms dominant and recessive can be used in referring to the alleles involved in expression or to the phenotype itself.

In addition to using phenotype to describe a particular trait, the term phenotype can also be used loosely to refer to the entire organism. Under this broader usage, though, the terms dominant and recessive are generally meaningless. And it can get more complicated. Although some traits, such as eye color, may not be influenced by environmental conditions, other traits can be heavily influenced (like our cake). For example, weight at one year of age is influenced by genes involved in the growth of an organism. Such traits are the basis of agricultural genetics and the large increase in the world’s ability to produce food that has taken place over the last century. Traits like weight at age, however, also depend on the feeding regime and other environmental factors. As a result, a phenotype can depend on both genotype and environment. We will consider these ideas in the next chapter.

Mendel’s laws and gamete formation

In contrast to mitosis, sexual reproduction results when a male and a female gamete—a sperm and an egg—unite to form a diploid zygote. Sperm and eggs are haploid—they carry only a single copy of each chromosome. A zygote carries the same amount of genetic information as the parents did because each gamete carried one haploid set, and when the gametes united, the diploid complement was restored. Gametes are produced by meiosis, a process similar to mitosis, but meiosis requires two divisions (Figure 4). One of the divisions separates each chromosome product of a homologous pair and the second division generates products that have only a single copy of each chromosome. The Noah’s ark analogy can be extended here if you think of arks as cells and the chromosomes as animals. If each pair of animals on two different arks produced one offspring, a third ark could be populated by combining the offspring from the two arks (Figure 5).

The number of products that results from meiosis usually differs between males and females. In producing sperm, each meiotic process yields four spermatozoa. In producing eggs, the chromosomes are apportioned in the same way. But in most animals only one product ultimately becomes an egg—the other three potential products are “sacrificed” to ensure that the single egg has sufficient yolk.

Mendel also considered the results of inheritance of one, two, and more distinct traits. Although Mendel knew nothing about chromosomes, he deduced the behavior of the two genes for each trait, and the behavior is exactly described by the way in which chromosomes are distributed during meiosis. For a single gene pair, the gametes carry only one of the two chromosome pairs. The diploid complement is restored at fertilization (Figure 6). If two pairs of genes are carried by different chromosomes, they will behave independently. That is, inheritance of one trait is in no way connected with inheritance of the second trait on the other chromosome. The idea is similar to tossing a coin and rolling a die. Whether the coin lands with heads or tails up has nothing to do with which face of the die is up. The easiest way to visualize this process is to use a matrix to follow the possible results after two generations of matings (Figure 7).

Sexual reproduction and diploidy provide a mechanism that generates huge amounts of genetic variation. For example, pink salmon have 26 pairs of chromosomes. If each chromosome pair carried only a single heterozygous locus (and of course each usually has many, many more, some exceeding 1,000), $2^{26}$ different (haploid) gametes could be produced from the combination of alleles at the different loci. The number of diploid individuals that could be generated is $(2^{26}) \times (2^{26})$. That is more than 67 million different gametes and more than 4 quadrillion (15 zeros) genotypes. Those numbers result from just 26 variable loci. In reality, we will never see all of the genotypes that are possible because we will never see 4 quadrillion pink salmon.
Meiosis is the process that exactly divides the diploid set of chromosomes (and genes) into haploid gametes. When two gametes unite, the diploid sets are restored. Most of the cellular events for the meiotic phases shown are the same as in mitosis.
How do we explain dominance?

The phenotypes produced by alleles at a locus usually result from the expression of the information carried in the instructions. That is, the cake that results from the recipe. Some alleles have faulty information—for example, if the recipe instructed you to clean up and discard everything before the cake batter would be poured into the pan and baked, no cake would be produced by that faulty recipe. If you had two recipes, however, and one recipe was complete, you would still successfully produce a cake if you followed both recipes, and the result (phenotype) would be “successful cake.” If both were copies of the faulty recipe, however, the phenotype would be “no cake.” Many recessive alleles carry defective information. Albinism results when an individual carries two defective alleles that specify one of the enzymes involved in the biosynthetic pathway for the pigment melanin. The recessive wrinkled peas that Mendel studied resulted from defective instructions for an enzyme that synthesized starch. Starchy peas retain water and remain plump; without the starch, the peas shrivel and are wrinkled. Most recessive metabolic diseases (like cystic fibrosis and phenylketonuria) also result from defective instructions for key enzymes.

Codominant and partial dominant phenotypes

Even though Mendel’s laws are most often explained from traits that are expressed as dominant or recessive phenotypes, the expression of all loci is not restricted to those modes. At some loci, both alleles are expressed; this is referred to as codominant expression. This is like the effect of wearing bowling shoes—ordinarily, one shoe is designed for sliding (the left shoe for right-handed bowlers) and the other for breaking. The function of both shoes is expressed in the complete phenotype of the bowling action. The standard example for codominance is the blood type AB, the expression of which results from the chemical groups attached to the outside of blood cells. The group attached as a result of expression of the A allele differs from that attached as a result of expression of the B allele. Type A blood can result from a genotype homozygous for the A allele (AA) or from a genotype that involves another allele at that locus, the o allele.
Figure 6. Mendel’s first law. The two alleles at a locus separate during gamete formation. The diploid state is restored at fertilization. Note that a recessive trait (albinism here) is not expressed unless it is homozygous. In the $F_1$ generation, both males and females are heterozygous. The products in the second generation are shown in the grid as the intersection between possible sperm and eggs. Second generation individuals will be expected to have a phenotypic ratio of 3 dominant: 1 recessive. Note that there are two ways to obtain heterozygotes in the second generation—the sperm can contribute the $A$ and the egg the $a$ or vice versa.
What Does Genetics Have to Do with It?

Figure 7. Mendel’s second law. Inheritance of alleles at one locus is not influenced by the inheritance of alleles at other loci. Here we follow two traits, one for coloration—pigmented (A) or albino (a)—and a second trait for tail spotting—no spotting (B) is dominant to spotting (b). Gametes that have four different composite genotypes (AB, Ab, aB, and ab) can be produced by the F1 double heterozygotes (AaBb). The grid shows all possible ways in which the different possible eggs and sperm could combine. Note that in the second generation of these crosses, the phenotypic ratio is 9 that are dominant for both traits: 3 dominant for the first and recessive for the second: 3 recessive for the first and dominant for the second: 1 recessive for both. Of course these frequencies will only be accurate if many progeny are produced because the inheritance of alleles follows laws of probability.
which is recessive and attaches no chemical group. The genotype \textit{oo} produces type O blood, in which neither of the chemical groups is attached to the red blood cell. Similarly, type B blood can result from either \textit{BB} or \textit{Bo} genotypes.

Another mode of inheritance is partial dominance. An example of this is carnation color. Red carnations are homozygous for the \textit{R} allele (\textit{RR}), white carnations are homozygous for the \textit{r} allele (\textit{rr}), and pink carnations are heterozygous (\textit{Rr}). Expression of the \textit{R} allele apparently produces red pigment, but the amount produced by a single \textit{R} allele is insufficient to turn the carnation red; red requires the combined expression of both \textit{R} alleles.

\textbf{Sex-linked and nuclear genes versus mitochondrial genes}

Most organisms also carry some genes that are not diploid and do not obey Mendel’s laws. In some species, the chromosomes that determine the sex of an individual are a special pair of chromosomes. Just like all of the normal chromosomes, sex chromosomes pair during meiosis and are sorted. However, the two sex chromosomes often do not carry the same set of genes (loci) in many species (including humans). As a result, an individual may carry only a single copy of a gene (allele) on a sex chromosome. For example, human females have two X chromosomes; and human males have just one X chromosome but also carry one Y chromosome. The human Y chromosome carries very few genes, so males carry only a single allele at each locus (one copy of each gene) that is on their X chromosome. And they inherited that allele from their mothers. Traits that are coded for by genes that are on the X chromosome are referred to as \textit{sex-linked} traits. Two sex-linked traits in humans are hemophilia (the gene specifies instructions for a blood clotting factor) and color blindness. If a male inherits a defective gene for one of these traits, he will be a hemophiliac or be color blind because he has only that one copy of the allele, the one on the single X chromosome he received from his mother. The incidence of sex-linked phenotypes is higher in human males than in females. Relatively few genes are sex-linked; geneticists have little specific knowledge about sex-linked traits in fishes.

Another set of genes that are not diploid is carried by mitochondria instead of the nucleus. Mitochondria are the power plants of the cell. They are structures inside the cell (subcellular) and are responsible for converting chemical energy into a form that is useful to the cell (e.g., adenosine triphosphate—ATP, a high energy molecule, which is the intracellular energy currency). Mitochondria carry their own instructions for some, but not all, of their protein components—they have their own DNA for those proteins. Two mitochondria are produced as a result of division by a single mitochondrion. In vertebrates, mitochondria are passed to offspring by the mother, not the father. Consequently, mitochondrial molecules are haploid (single copy) and clones (identical in offspring from the same mother). We will examine the mitochondrial genome in more detail in Chapter 11 (Is a Rougheye Rockfish Really a Rougheye Rockfish?).

\textbf{Summary}

In this chapter we examined the basic rules of genetics. The information that defines an individual or species (phenotype) is encoded in genes, which are arranged on chromosomes. Mendel’s laws describe the cellular process of meiosis, which separates the two alleles carried by diploid organisms to produce haploid gametes, and subsequent fertilization (sexual reproduction), which restores the diploid number in offspring. Because the two alleles carried by an individual can carry slightly different versions of the instructions, sexual reproduction can generate enormous genetic variability.
This is the second in a series of articles that considers the role genetics plays in conservation and management of our fisheries resources. In the first chapter, we examined the basic genetics of inheritance, Mendel’s laws. In this chapter we look at the genetics of populations, because most of the genetics applications that pertain to conservation and management of fish populations involve the behavior of genes (alleles) in a population and the genetic differences among populations. Of course, the inheritance of alleles and their individual expression follow Mendel’s laws, but now the population is the focus of our interest, rather than progeny from specific crosses. In addition, many of the traits important in survival and adaptation to wild environments, and to improvements in aquacultural applications, result from the combined expression of multiple loci. Traits such as size, fecundity, and thermal tolerance do not result from the expression of single loci, and their study requires some additional approaches.
**Gene pool**

The term gene pool refers to the total aggregation of genes in a population. A gene pool envisions a population as a set of haploid gametes (sperm and eggs) that can unite (fertilization) at random to form diploid individuals. The gene pool concept assumes that random mating occurs in the population. However, theoretical studies have shown that the random mating assumption does not need to be rigidly adhered to and that a gene pool is a very useful simplification, which accurately describes the genetic composition and genetic dynamics of a population for most purposes. In the gene pool of a population, we look at every individual and count the types of alleles that exist at a locus. For simplicity, we examine one locus at a time. We quantify and describe populations in terms of their allele frequencies, which are the relative proportions of each type of allele at a locus. For example, allele a may have a frequency of 0.3 and allele A has a frequency of 0.7. These are referred to as allele frequencies. The allele frequencies at a locus add to one. Because it is impractical to examine every individual in most populations, allele frequencies are usually estimated by taking a sample of individuals from a population and counting the numbers of each type of allele that they carry. Larger samples produce more accurate estimates of allele frequencies.

We will see below that most population genetics processes act by directly altering allele frequencies. For those processes, the gene pool concept works well. We will also see that the gene pool concept may not work simply for quantifying selection, because selection acts on phenotypes and alters phenotypic frequencies, which in turn changes allele frequencies indirectly.

**Sources of variation and genetic change in populations**

The frequencies of alleles at a locus can be altered by several processes. The common processes, sometimes referred to as “forces,” are mutation, selection, gene flow (= migration), and random drift. A simple definition of evolution is a change in allele frequencies in a population. The change can result from the action of any of those forces individually or together.
How Genes Vary in Fish Populations

15

Most existing organisms reflect the product of eons of genetic processing. Many of the fish populations that we see today are the results of tens of thousands of years of response to a sequence of local environments. Pacific salmon are especially interesting because their evolutionary history is strongly related to the glacial history of the area in which they are now found. Pacific salmon are anadromous—they spawn in fresh water, emigrate to the marine environment where they feed and grow, and return surprisingly accurately to their natal streams to spawn. When glaciers advance, many streams become covered with ice and are no longer available to salmon.

Geologic time is appropriate for considering the emergence of species and for examining the postglacial colonization of anadromous species and even for the recent history of many marine species. The times are in thousands to millions of years. In contrast, the actions of most conservation and management biologists are scheduled for time frames of less than 100 years and often a manager’s career of about 20 years. Their focus on such short times does not mean that resource biologists are unconcerned about longer times. But, from a practical perspective, their actions require evaluation over these much shorter times. One view of the success that a manager has had during a career is that the resource did not decline during his or her stewardship.

Harvest decisions can exact much higher mortality levels than naturally occurring mortality levels. The consequences of depletion for long-lived species, such as the rockfishes, can be disastrous. Many rockfish species live more than 100 years and do not mature until their teens or twenties, just like us. Consequently, their life history timing is roughly the same as ours or longer. As Milton Love (University of California Santa Barbara) points out, a large yelloweye rockfish could have voted for Abraham Lincoln. Actions taken during a manager’s career often cannot be thoroughly evaluated in that time frame; depletion often can be detected, but the results of restoration actions are difficult to detect.

Sidebar 1

TWO TIME FRAMES: GEOLOGIC AND MANAGER’S

Most existing organisms reflect the product of eons of genetic processing. Many of the fish populations that we see today are the results of tens of thousands of years of response to a sequence of local environments. Pacific salmon are especially interesting because their evolutionary history is strongly related to the glacial history of the area in which they are now found. Pacific salmon are anadromous—they spawn in fresh water, emigrate to the marine environment where they feed and grow, and return surprisingly accurately to their natal streams to spawn. When glaciers advance, many streams become covered with ice and are no longer available to salmon.

Geologic time is appropriate for considering the emergence of species and for examining the postglacial colonization of anadromous species and even for the recent history of many marine species. The times are in thousands to millions of years. In contrast, the actions of most conservation and management biologists are scheduled for time frames of less than 100 years and often a manager’s career of about 20 years. Their focus on such short times does not mean that resource biologists are unconcerned about longer times. But, from a practical perspective, their actions require evaluation over these much shorter times. One view of the success that a manager has had during a career is that the resource did not decline during his or her stewardship.

Gene flow

Gene flow, or migration, describes the movement of genes from one population to another (Figure 2). In salmon, gene flow results from strays that do not home accurately, whether by accident or design. However, only the strays that successfully produce offspring in the new stream contribute to that population. What this means is that just because straying has been documented does not mean that gene flow has occurred, because genes from the stray may not have been passed to the progeny. There is evidence that in some instances the genes of translocated fish do not enter the recipient gene pool in substantial numbers, but it would be folly to count on such failures. The effects of gene flow can accrue very rapidly. If translocated fish can successfully interbreed with native fish, in some situations their genes can swamp the native population in a very short time. Is this an issue? Stay tuned for other chapters.

Selection

Most people who have been exposed to Darwin’s ideas on natural selection think of selection as an important force, if not the governing force. Note, however, that the genetic variation must exist (mutation) before selection can play a role. Selection is any force that alters the genetic composition of a population by differentially influencing reproduction or survival. Warm water could reduce the fertility of some individuals more than others. In that instance, temperature would be the selective force.

When one thinks about selection, it is important to realize that an individual either contributes to the next generation or it doesn’t (e.g., genetic life or death). Evolutionary or genetic fitness (more commonly referred to as just plain fitness) is a measure of an individual’s contribution of genes to the next generation. When we say individual, we actually mean the particular phenotype. If the
What Does Genetics Have to Do with It?

biggest, toughest, meanest son-of-a-gun around has an accident that makes him sterile before puberty, his fitness is zero. Consequently, it is populations and not individuals that evolve; populations not individuals experience changes in their allele frequencies in response to selection pressures.

Selection acts on phenotypes and sometimes it will favor or select against both AA and Aa genotypes (A is dominant and both genotypes have the same phenotype relative to fitness), compared to the aa genotype (Figure 3). The connection between altering phenotypic frequencies (which are based on genotypes of two alleles) and allelic frequencies is not simple, and the gene pool concept does not directly apply. One of the reasons low levels of some genetic diseases, like phenylketonuria (PKU), persist in the human population is that the recessive alleles responsible get weeded out only in homozygotes (aa). Heterozygous individuals (Aa) pass both alleles to the next generation because their phenotype is unaffected, even though they carry the a allele.

The elements of the natural environment, such as stream temperatures or flows, can change according to year-to-year fluctuations or global changes; as the environment fluctuates or changes, the way in which populations respond genetically also changes. Some species, such as Pacific salmon, experience a diversity of environments during their lives: in the gravel as embryos and newly hatched individuals, possibly in freshwater streams or lakes as young juveniles, in estuaries as young fish, and on the high seas as growing fish. A particular allele may behave differently in different environments, but it is very difficult to measure the effects of different segments of the life history of a fish or population. This is because we cannot usually monitor them throughout their lives—we only see them early in their lives and not again until they return to spawn.

Selection can occur either inadvertently or purposefully in culture situations. Artificial selection—breeding—is the primary reason that agriculture developed so rapidly during the last
How Genes Vary in Fish Populations

Strong selection (which causes large changes in allele frequencies) can change allele frequencies in very short times. In contrast, weak selection (only small changes in allele frequencies) may be scarcely noticeable, and even modest levels of selection can be very difficult to detect in a manager’s time frame.

Random drift
Random drift is the term used to describe how the gene pool of a population changes from generation to generation. As its name indicates, random drift is a random process. Random drift occurs because the genes that are passed between generations are a subset of the parental gene pool—winners of the genetics lottery—which is often sampled nearly at random;

and only the complete set of alleles in the population represents the gene pool without error. The gametes that successfully produce the next generation can be viewed as a sample of alleles from the gene pool. Because the sample of alleles does not include every gene in the gene pool, the offspring may not perfectly represent the gene pool of the previous generation. For example, if we flip a fair coin, we expect 50% of the tosses to be heads and 50% to be tails. If we tossed the coin 1,000 times, on average we would see close to 500 heads, but if we tossed it only 10 times, we might reasonably expect to see any number from 3 to 7 heads. The same process occurs in the reproduction process, during which alleles are sampled from the population. Because samples may not perfectly represent their source, the allele frequencies in small populations vary randomly between generations. Over a sufficiently long period, small populations may actually lose genetic variability as a result of repeatedly occurring random drift. A short-term dip in population size can alter allele frequencies for a long time (in a manager’s time frame) even if the population size expands. These dips in number are referred to as bottlenecks (Figure 4).

If a very small number of individuals initiates a new population, for example in a hatchery, the allele frequencies may differ substantially from the donor population. This is referred to as a founder effect. In Lancaster County, Pennsylvania, the Amish, a religious sect that colonized North America more than 200 years ago, have a high incidence of a recessive disorder called six-fingered dwarfism. The prevalence of the disease is due in part to the small number of individuals that founded the population. The persistence results because there is little gene flow into the population.

Biochemical genetic markers versus quantitative or metric traits
The focus of most population genetics studies is the genetic variation that occurs within and between populations. In fact, the biological function of loci at which the variation exists is often irrelevant to geneticists because they are using the differences observed as indices or markers to evaluate genetic structure. In fact, the variation at many of the loci
Figure 4. Random drift generates random fluctuations in allele frequencies. In this figure, the possible effects of a one-time bottleneck are illustrated. The direction of change is unpredictable. However, the average magnitude of change is related to the number of individuals that exist when the population is at its minimum size. Severe population declines usually cause more severe changes than moderate declines. Loss of relatively uncommon alleles is symptomatic of population declines, whether abrupt and severe or mild but chronic.
that geneticists use is not significantly influenced by natural selection. Such variation is referred to as **neutral** variation. Most population geneticists study biochemical traits that can be easily detected. Alleles of these loci in aggregate often serve as **markers** (like natural tattoos) that can be used to identify populations. Be careful, though—geneticists have not yet discovered markers that can unequivocally track the natal origin of each single fish.

Although several different methods are used to obtain data, most of the methods directly or indirectly detect differences in the DNA sequence (the actual instructions in the blueprints) of target genes or of the protein products that those genes specify. What makes them target genes? Either through happenstance or deliberate searching, geneticists have learned that the target genes have more than one detectable form (i.e., allele). Recall that the ABO blood group has three alleles (A, B, and O), and some loci may have hundreds of alleles. Of course each individual can carry only two of the many available alleles if they are diploid, but the population can have many. For example at locus A, if Jim has an A₁A₂ genotype, Mary has an A₃A₄ genotype, and Lou has an A₅A₆ genotype, we can count five different alleles (A₁, A₂, A₃, A₄, and A₅) at locus A, which are found in just those three individuals. Note that because they are diploid, each individual carries exactly two alleles. The rest of the population may possess even more alleles. Genes that are targeted for study and are in the nucleus ordinarily obey Mendel’s laws and may be referred to as simple Mendelian traits. We will consider these molecular tools in a later chapter.

Even though allele frequencies at most of the loci that are mined for population genetics data are not influenced substantially by natural selection, deleterious genes do exist. Two examples in humans are deleterious genes that cause phenylketonuria (PKU—the inability to metabolize the amino acid phenylalanine) and sickle cell anemia. We see in Figure 2 that afflictions resulting from recessive alleles may occur if an individual carries two abnormal copies of a gene. For these recessive traits, afflicted individuals can have either one or two normal copies of the instructions, but only individuals homozygous for the defective gene are affected. The normal phenotype is dominant. It is interesting to note that the severity of PKU can be reduced by altering the environment, in this case by eliminating phenylalanine from the diet.

Most population geneticists study simple biochemical genetic traits, which are generally not the traits that are important in fish husbandry or, for the most part, to a population adapting to a local environment. In both fish husbandry and adaptation, traits such as growth rate, fecundity, and temperature tolerance are critical. Clearly, such traits are the net result of the expression of alleles at multiple loci as well as the environment in which they are expressed. Unfortunately, it is usually not possible to disentangle the genes for these quantitative or polygenic traits and study them one at a time (see sidebar 2).

**Summary**

The central concept that will be important to us in the following chapters is local adaptation, which is the process by which (the gene pool of) a population evolves to be most productive under the environmental conditions experienced by the population. Different populations (actually their gene pools) can differ because adaptations to different environments (locations) often require different genetic solutions. We briefly described the primary forces that alter the gene pool: mutation, gene flow, selection, and random drift. Finally, we learned that although each gene follows Mendel’s laws, the aggregation of genetic information that is present for many of the traits, particularly the traits that are acted on by local adaptation, is the result of the expression of multiple genes (quantitative traits) and must be dealt with in a different way than simple Mendelian traits.
What Does Genetics Have to Do with It?

Quantitative traits result from the combined action of several to many genes. Such traits are also called polygenic or metric traits. Many of the traits important for local adaptation, such as salinity or temperature tolerance and migration timing, are influenced by the alleles at many different loci. Even though inheritance at each of the contributing loci obeys Mendel’s laws, the influence of individual genes usually cannot be separated from the total. Consequently, we have to resort to statistical methods to dissect the role of genetics in the expression of these complicated traits. Let’s see how that is accomplished.

The problem is that an individual’s genetic instructions are not the only contributor to its phenotype. If we consider a trait such as weight at age, it is clear that the feeding regimen and many other environmental factors play a role in the expression of the trait. In fact we can look at the environment as a black box that filters the genetic instructions (genotype) in producing the phenotype (Figure 5).

Figure 6 shows that the individuals in a population have a distribution of phenotypes (see Figure 5 also), which results from the expression of the genotype in the population’s environmental background. Some of the variation that we see in a phenotype, such as weight, derives from genetic differences between individuals, and some of it is a result of the different environments that they experience. Even identical twins do not experience exactly the same environments. A humped curve is ordinarily what you will see if you plot the abundances of different quantitative traits (such as size) in a population. Most of the individuals are near the middle; extraordinarily large or small (extreme) individuals are relatively rare and at the ends of the distribution. The midpoint of the distribution is the mean. The extent of the spread of phenotypes, which reflects the variation among individuals, is quantified as the variance. We cannot conveniently graph both the mean and variance in the same figure, but we can show the standard deviation, which is the square root of the variance (S.D. = \(\sqrt{\text{variance}}\)) (Figure 6).

The goal of quantitative geneticists is to partition (separate into contributing components) the variation of the phenotypic trait. That is, separate the variance of the phenotype (or \(V_p\) in statistical terms) into the genetic variance (\(V_g\)) and environmental variance (\(V_e\)) components:

\[
V_p = V_g + V_e
\]

When we described Mendelian inheritance, we did so in terms of dominant and recessive traits. For many traits, however, both alleles at a locus contribute to a phenotype, although not necessarily equally. We provided carnation color (red, pink, and white) as an example of incomplete dominance. For a quantitative trait, such as size, a homozygote (AA) might have a value of 2 pounds, the other homozygote (aa) might have a value of 1 pound, and the heterozygote (Aa) might be intermediate at 1.5 pounds. In this example, the effects of the alleles are additive; the A allele contributes 1 pound toward the phenotype and the a allele contributes ½ pound. In addition, alleles at different loci may act synergistically to produce an unexpected result. This is referred to as epistasis. Often the total genetic variation (\(V'_g\)) can be further partitioned into one component that quantifies variation due to
additive effects ($V_A$) and another that quantifies the nonadditive effects ($V_N$), dominance and epistasis (Figure 7).

Geneticists have coined a term that quantifies the relative importance of heredity in expression of a phenotype in a population, the **heritability** of the trait. There are two different heritabilities. **Heritability in the broad sense** ($H^2 = V_G/V_P$) estimates the proportional influence of all genetic influences, and **heritability in the narrow sense** ($h^2 = V_A/V_P$) estimates the proportion of a phenotype that is attributable only to additive genetic influences. The latter is used to predict the results of selection in agricultural and aquacultural directed selection programs.

To study quantitative traits, geneticists conduct breeding experiments. Two different approaches may be used. In one approach, individuals are selected for their trait (say the largest 10%) and the trait is measured in their offspring (Figure 8). Of course, it is usually necessary to follow a nonselected control population, which serves as a reference for the **response to selection**, the name of this approach.

Another approach is to make multiple crosses among individuals that were chosen at random from the population. Heritabilities are deduced from these crosses by measuring the phenotypic similarities that related individuals share with each other but do not share with other members of the populations. The proportions of shared alleles are also factored into this computation. For example, large parents generally produce large offspring and small parents generally produce small offspring.

**Figure 7.** The phenotypic variation ($V_P$) of a quantitative trait in a population is the total of the variation (the statistical variance) that is observed for the trait. $V_P$ is the sum of genetic influences ($V_G$) and environmental influences ($V_E$). $V_G$ can be further separated into genetic contributions that reflect additive effects of alleles ($V_A$ = additive variance) and contributions that result from interactions between genes and loci ($V_N$ = nonadditive variance).

**Figure 8.** Response to selection. Individuals in one tail of a phenotypic distribution are used as breeders—e.g., the largest ones. This tail, which is the parents, has a mean of $M_{Par}$. If their offspring have an average phenotype ($M_{Off}$) that exceeds the population mean ($M_{Pop}$)—in size for this example—there is a genetic component for the phenotype. The bigger the deviation from the population mean, the greater the genetic component. Heritability can be estimated: $h^2 = (M_{Off} - M_{Pop})/(M_{Par} - M_{Pop})$. 
The trait "size" is heritable and the heritability \( (h^2, \text{strength of genetic contribution}) \) can be quantified in the terms shown in Figure 7. Offspring share half of their alleles with each parent (from the parent's gamete). For many traits, the progeny must be tracked to adulthood or marked by family and recovered at maturity in order to obtain comparable measurements of relatives. Breeding experiments are expensive and usually require dedicated facilities, which make them difficult to conduct on wild populations. However, quantitative traits are as a rule far more important to an individual's and population's success than the vast majority of biochemical genetic traits used by population geneticists to study population structure. Two other problems are that heritabilities are population-specific characteristics and when they are estimated, they are usually estimated for a particular environment or generation. Finally, \( h^2 \) estimates the relative proportion of additive genetic variation, but it does not reveal the actual allelic composition.

The similarity (statistical correlation) between parents and offspring can be demonstrated by a plot of an offspring's value (e.g., length) versus its parent's value (Figure 9). Estimates of the average similarities can be made with regression analysis, a statistical tool that quantifies the relationships between the sizes of parents and offspring. If a trait is heritable, one would also expect to see a relationship between the sizes of siblings, cousins, half-siblings, and so on for any pair of “blood” relatives. Other statistical methods are used to quantify these relationships and extract information about the heritability of traits.

**Figure 9.** This is a plot of the length of a son on the y-axis against the length of its father (= sire) on the x-axis. There is a linear relationship, although it is not perfect, which indicates that there is a genetic basis to length. If you were expecting the pitch of the line (slope) to be one (an incremental increase in sire size produces the same incremental increase in the size of its son), you would have been wrong. First, only half of the genetic information for size is inherited from the father, the mother provides the other half; and second, environment also contributes to the variation in size of both fathers and sons.
This is the third chapter of our journey into the realm of genetics and its role in fisheries conservation and management. In the first two chapters, we examined elementary genetics concepts and the jargon that is used to convey those concepts (1. Even Fish Obey Mendel’s Laws, and 2. How Genes Vary in Fish Populations). We saw how mechanisms that transmit chromosomes during cell division—mitosis and meiosis—led to Mendel’s laws, which predict the basic nature of inheritance. Meiosis is also important because it shuffles genetic variation to generate enormous variety in living things. We also said that an important key to the role of genetics in conservation and management is the differences between populations produced by local adaptation. In this chapter, we examine a salmon population to learn about the historic events that have influenced it. That history is tremendously important to understanding the population, its genetics, and its relationship to other populations.
During the Pleistocene epoch, which extends back 2 million years, there have been periodic changes in climate, sea level, glacial extent, and oceanographic conditions. It is important to realize that our present-day conditions, which followed the last glacial advance (and presumably precede the next) are an extreme. We now enjoy a much warmer climate than existed during most of the Pleistocene and, as a result of greenhouse gases, the earth will probably warm up even more. The variations in conditions have been deduced from the global ice and seawater volumes that are recorded in the chemistry of marine sediments and ice cores from Greenland and Antarctica and from geologic evidence that lower sea levels, more extensive glaciation, colder sea surface temperatures, and more extensive and southerly (northerly in the Southern Hemisphere) sea ice were typical during glacial advances (sidebar 1). Warm interglacial periods, such as we are now experiencing, have occurred about every 100,000 years or so, but, again, most of the Pleistocene was considerably colder.

Salmon populations along the present-day Sea of Okhotsk, Bering Sea, and Gulf of Alaska are separated geographically by landmasses and oceanographically by the different currents that flow into or between the oceanic basins and influence marine migration routes. Because of the separation, relatively few individuals stray between these geographic regions; and with no gene flow tying the areas together, genetic divergence can develop. The extent of geographic separation was greatly influenced by the limits of glacial advances and by the changes in climate. In their northern range, salmon populations experienced increased isolation between the marine basins as a result of lower sea level, loss of freshwater habitat to increased ice cover, and more extensive sea ice. Just as recent fluctuations in salmon productivity have resulted from comparatively minor climate changes, less favorable freshwater and marine environmental conditions undoubtedly decreased the sizes and numbers of populations substantially, further isolated the populations that remained in these regions, and probably increased the genetic divergence. Let’s look at this in a bit more detail.

At the last glacial maximum (LGM), which was about 14,000-20,000 years ago, the paleogeography of the North Pacific coastline was severely altered; both the sizes of the enclosed basins as well as the circulation among them were drastically reduced. Marine sediment records (sidebar 1) suggest that during most of the last 100,000 years sea level was 50 meters lower than at present, and at the LGM the level was 120 to 130 meters lower because much of the earth’s water was tied up in ice sheets. The lowered sea level exposed much of the shelf in the Bering Sea down to the continental slope. The resulting landmass, Beringia, broadly connected western Alaska and northeastern Russia.

At the LGM, Sakhalin Island was connected to the Asian mainland and to Hokkaido Island, which blocked the northern outflow of the Sea of Japan. Also, the connection between a smaller Sea of Okhotsk with the Pacific Ocean through the Kurile Islands was restricted. In the Bering Sea, the extensive continental shelf was exposed, and blocked circulation through the Bering Strait (Figure 1). To the south, the Aleutian Islands and the Alaska Peninsula were joined to about 170°W, and many of the islands to the west were also connected, which dramatically limited water movements between the Gulf of Alaska and the Bering Sea.

Coastal areas of the Gulf of Alaska, including most of the continental shelf, were extensively glaciated, although a few isolated areas of the outer coast were probably ice free. The ice cover of the eastern Aleutians coalesced with the Alaska Peninsula, and ice caps covered all the major islands. In these areas there may have been ephemeral streams from melted snow or ice at the southern and western margin of the Gulf of Alaska near the present continental shelf break. However, salmon-spawning habitat along the Gulf of Alaska was almost completely eliminated at the LGM and much of the Bering Sea coastline was probably icebound most of the year, which reduced available freshwater habitat.

Obviously, salmon did not completely disappear during the glacial advances. However, their ranges were radically reduced because streams were covered with ice. The salmon that survived did so in several regions that escaped glaciation. There were several locations, termed refugia (just one is a refugium), that continued to harbor salmon, although the abundances of salmon in most of those areas declined substantially.
One of the major areas in which salmon survived was to the south of the major glaciation in the western United States—the Columbian refugium. A second area was the Beringian refugium, which was then the northern boundary of the Bering Sea (Figure 1).

In addition, some populations probably survived in several different areas in Asia. Geologic evidence suggests that the LGM did not affect Asian streams as broadly or severely as Alaska streams. Glaciation was much less extensive and included some alpine glaciation, but few glaciers extended to tidewater, except possibly on the southeast side of the Kamchatka Peninsula. Freshwater habitat was probably not reduced to the extent that it was along the Gulf of Alaska coast. Nevertheless, the Sea of Okhotsk, like the Bering Sea, probably had sea ice cover much of the year as it does now in the winter. It is also likely that the Sea of Japan was landlocked during the peak of glaciation.

In addition to the larger refugia, there is evidence that some unglaciated areas persisted through the LGM of the Pleistocene that could have been very small refugia for salmon. Such areas probably included the Queen Charlotte Islands in Canada and possibly the outer coast of Prince of Wales and Baranof islands in Southeast Alaska.

The much lower sea level (more than 300 feet) exposed the continental shelf of the Bering Sea. As mentioned, the exposed land and adjacent regions are called Beringia. Its southern edge followed what

Figure 1. Approximate coastline and extent of glaciation at the last glacial maximum. The dark green area shows the modern coastline, the light green area is land that was exposed by the drop in sea level, and the enclosed black areas were ice masses. Note that most of northern North America to the east (not shown) was ice covered at this time.
What Does Genetics Have to Do with It?

Foraminiferans (Figure 2) are small microscopic animals that live in all the world’s oceans at virtually all depths. They construct “skeletons” from insoluble carbonate salts of calcium and other metals. When pelagic species (those that live out in the open ocean near the surface) die, they settle out on the ocean floor and their skeletons help create layers of sediment. In many places in the oceans, there are no currents or other forces to disturb the sediments. In those places, the deposits of the foraminiferan debris generate a sequential record of the minerals that were in the ocean at the time they were alive. The older layers lie beneath the more recent layers, so as we look deeper and deeper into the seafloor, we encounter older and older remains. By very carefully extracting a long core, paleontologists can look back in time at the carbonate that has been deposited. The chemical compositions of the skeletons resemble the chemistry of the water that surrounded them when they were made. The chemistry of the water differs between the times when there was substantial glaciation and times when there wasn’t.

What is carbonate? Why would one want to look at carbonate? Carbonate is made of oxygen and carbon, which are components of seawater. Water (H₂O) and carbon dioxide (CO₂) combine in the sea to form carbonic acid (H₂CO₃). We are familiar with carbonate in the form of baking soda—sodium bicarbonate (NaHCO₃), which reflects the substitution of hydrogen (H) in carbonic acid by sodium (Na). Recall that adding an acid (like vinegar) to baking soda causes the release of carbon dioxide gas. Clamshells are largely made of calcium carbonate (CaCO₃; the crystalline structure is one of several carbonate minerals), which is insoluble in many areas of the world’s oceans. One of the special components for dating the sediment is the oxygen (O) in the carbonate. In the ocean, the oxygen atoms in water can exchange with the oxygen atoms in the carbonates, until the insoluble carbonate salts form. In the insoluble carbonate salts, however, the oxygen atom can be tied up for very long times (millions of years).
It turns out that there are several naturally occurring kinds of oxygen atoms. One has 8 protons and 8 neutrons (16O; about 99.7% of the earth’s total oxygen) and the next most common type has 8 protons and 10 neutrons (18O; about 0.2% of the earth’s total oxygen). Compounds like water (H₂O) that have the 18O are slightly denser than ones that have the 16O. One of the consequences is that water molecules that have 18O do not evaporate quite as readily as ones that have 16O. When the climate is cold, much of the evaporated seawater is deposited as snow and ultimately as ice in glaciers. Amazingly, enough 16O is differentially removed from the seawater that differences in abundances of 16O and 18O can be detected. The relative abundances are quantified by a term called δ18O (delta 18 oxygen), which is an index for the relative amount of water (which primarily carries 16O) that is tied up in glacial masses at a particular time. The skeletons of the foraminiferans incorporate the available oxygen types at that particular time. So, by examining slices of the sediment core, we can determine the history of glaciers.

Water that evaporated from the ocean was deposited on the land, which increased the sizes of the glaciers. At the last glacial maximum, sea level was about 120 meters lower, more than the length of a football field, than it is today. That means that many of the shallow areas in the ocean were exposed. Other geological evidence corroborates the sea level drop. Figure 3 plots the δ18O index of sequential slices from a core taken from the deep ocean. The curve reflects warmer periods between glacial advances (called interglacial periods) at the top, and cold periods of glacial advances near the bottom. At the right is an indication of the approximate sea level that would correspond. Other cores, such as ice cores from Antarctica and Greenland, have shown very similar profiles. Note that during the vast majority of the time over the last 400,000 years, the earth had much more ice and was much colder than today. The last interglacial period was about 125,000 years ago.
is now the continental shelf break and extended to Russia, forming the Bering land bridge. This area had few glaciers because it was arid; most of the moisture in the atmosphere was extracted by mountains to the east and south. The Bering Sea was much smaller during this time and appears to have had sea ice cover much of the year. Seasonal sea ice may have persisted as far south as 54°N, nearly to the western Aleutians and Commander Islands, for 6-8 months a year during the LGM. Mouths of drainages from interior Alaska were displaced far from their present locations; there were flows that supported some salmon populations, although the freshwater environments were less favorable than at present.

Harsh conditions greatly reduced biological productivity of marine surface waters over the entire region. Microfossils indicate that the subarctic Pacific Ocean was similar to the present day Sea of Okhotsk, with cold fresh surface water and a highly stratified water column. Sea surface temperatures were probably 2° to 4°C colder than present throughout the year over most of the area. Off Japan, temperatures were even colder (6°C or more) at the LGM, which indicates that oceanic current patterns differed from today's.

As we trace the history of populations that emerged following the retreat of the glaciers, we illustrate the genetic factors that are important to population structure and how they arose. Even though we were not there at the time to record what happened, we know enough from geologic evidence, some of which was recounted above, and the recent biology and genetics of several salmon species to piece together a reasonable story.

Following deglaciation, it is likely that relatively few systems provided the source for salmon recolonizations. In the eastern range where habitat was ice-covered, reestablishment of salmon probably depended on colonization from the Bering refugium or more southerly refugia. Many of the streams populated by salmon then and even today in Southeast Alaska are short coastal streams that persisted only for short times in the geologic time frame of tens of thousands of years. Glaciers did not retreat instantaneously to expose new salmon habitat. Measured in human lifetimes, the changes were very slow. We can get a perspective from the rate of recession of glaciers in Glacier Bay and elsewhere in Alaska that is going on today. Historically, recession occurred in fits and starts and included short-term advances. For example, in 1986 and 2002, the Hubbard Glacier, near Yakutat, advanced and closed the entrance to Russell Fiord. The tidal flow between the fiord and Disenchantment Bay was blocked at Gilbert Point, and the water levels in the newly formed lake rose to 83 feet above sea level in 1986 and 61 feet in 2002 (http://www.fs.fed.us/r10/tongass/forest_facts/faqs/hubbard.shtml). There is also evidence that the Taku River, which drains just to the south of Juneau, Alaska, was blocked by the advance of the Taku Glacier in the mid 1800s. That blockage may have caused the Taku River headwaters to back up into the Yukon River system, which drains into the Bering Sea more than 1,000 miles of coastline from Juneau! Many other local events can block salmon passage, such as landslides and droughts, which occur sporadically. The idea is that deglaciation was slow and that many of the streams have not been continuously available.

The harsh conditions that existed during the LGM and the uncertain conditions that followed for a long time created a situation in which many local populations repeatedly went extinct. An ability to exploit available spawning habitat rapidly would have been an advantage; many local extinctions were probably followed by recolonization and resulted in extinction/recolonization cycles that will continue in most regions in a geologic time frame. To optimize the use of available habitat under these uncertain conditions, it would be advantageous for populations to have a higher straying rate (not put all of their eggs in one basket, uh, stream). If a stream became unavailable, some of the straying fish would probably still find a place to spawn (and perpetuate their genes). In contrast, if a population discovered and inhabited a stable stream, it would be less advantageous to have a high straying rate. It is likely that homing/straying have genetic determinants that respond to natural selection (sidebar 2).

With that insight, it seems likely that new spawning habitat became available slowly and unpredictably and that new habitat was colonized by
We are all familiar with the terms fitness and natural selection. But as familiar as we are with the terms, we need to revisit the concepts to be sure that we consider them in a genetics and evolution perspective. Many people go to fitness centers to “buff up” and enhance their health. The “fitness” that they achieve is not the same as the fitness that we are defining. Our definition of fitness is very simple and considers only reproductive success. Strictly, genetic fitness is determined by the number of genes that are contributed to the next generation (and that have the potential to be passed to subsequent generations). While buffing up may enhance the reproductive potential of some individuals, this has little to do with genetic fitness. The quintessential example of physical fitness versus genetic fitness is the mule, which was bred to do heavy work. A mule results from crossing a male donkey and a female horse, and a mule is sterile. So in spite of its physical prowess, a mule has a genetic fitness of zero.

Now let’s see how genetic fitness and evolution (through natural selection) go together. Consider a trait, say emigration time of pink or chum salmon fry. Clearly, fry that emerge and enter an estuary too early will not find food and they will perish. Similarly, fry that enter the estuary too late will not find food and they will perish. Fry will thrive only during a limited window of time in the spring and the most successful will enter at the optimum time during this window. So, what regulates emigration timing? Many factors, which include timing of spawning, rate of development, rate of yolk use, and ability to detect environmental cues (e.g., photoperiod, freshwater temperatures and flows), signal the suitability of the marine habitat; and there are many others that we humans are not clever enough to know. Most of those traits have some level of genetic determination; that is, they result from instructions in the individual’s DNA sequences.

Now focus on migration timing as a trait. The genes of parents that produce the fry that enter the estuary at the best time will be relatively more abundant in the next generation than genes of parents that produced fry that had poor timing. In the next generation, there will be more fish that have that particular timing. If the best estuarine conditions occurred at exactly the same time each year, eventually only the genes that resulted in that timing would exist in the population. The change in genetic composition, that is, the increases in abundance of genes that promote fitness, is evolution. We can visualize it in a simple diagram that shows the change of the frequencies of a trait that contributes to the fitness of individuals in a population. The green areas are stacked bar graphs; the lower category on the left gives rise to higher abundance on the right:

If heredity contributes to the successful phenotype the population changes genetically — it evolves.

Of course, the “prime time” differs from year to year. As a result, in some years some genes will be favored and in other years others will be favored. What will happen, though, is that the genes that produce migration timing on the edges of or outside the “prime time” window will become rare and genes for timing nearer the center of the window will become more abundant. It should be clear that it is important that the population have some genetic variability for emigration timing because it is impossible to predict in advance precisely when the best time will be. That variability will also serve the population (and species) as the climate varies, whether due to cyclic or global changes. Variation within the population is a form of insurance.

Another consideration is that all of the traits that contribute to fitness must be balanced. For example, larger fry usually outperform smaller fry. However, a large fry that enters the estuary too early will not outperform a smaller fry that leaves at the right time. In some cases, genes may contribute to conflicting traits. Natural selection resolves those conflicts. As an aside, Darwin formulated his theory of evolution before Mendel had figured out how heredity took place; Darwin had no idea of the mechanism of heredity.
vagrants. If we look at the dynamics of present day salmon populations, we see that their numbers vary substantially over time, even within the resource manager’s time frame. We also see that in years of high abundance, fish may appear in streams that rarely had fish previously. This is quite apparent with pink salmon (*Oncorhynchus gorbuscha*) in Southeast Alaska, but undoubtedly it is true for other species and areas. This observation suggests that, in addition to the irregular appearance of new habitat, colonization may have also occurred in fits and starts that reflected occasional periods of high abundance, presumably during periods of favorable environmental conditions. It is likely that much of the newly available habitat was colonized during these periods. The combination of being colonized by vagrants that were likely to produce offspring with some tendency to stray, and the likelihood that newly colonized habitats in an area probably drew from the same gene pool, suggests that many populations in a newly colonized area were probably genetically similar and remained that way for an extended period of time. This is referred to as historic genetic similarity.

Over time, as stream systems became more stable, differences in environments between streams (e.g., stream temperatures, stream flows) would contribute more strongly to the natural selection forces molding the genetic compositions of the populations. Under these more stable conditions and in the face of these local selection regimes, selection for improved homing would also occur. However, within a geographic area, it is likely that many streams share some environmental similarities and that some genetic exchange would persist. Natural gene flow has two advantages that would tend to keep it going. First, over longer periods of time, it is likely that occasionally populations will disappear as a result of catastrophic events. A small level of straying enables recolonization and would maintain the productivity (genes) of the system as a whole. Second, small populations lose genetic variation randomly through random drift (recall Chapter 2. How Genes Vary in Fish Populations). A low level of gene flow reintroduces lost and other variation. Population systems that are connected by low levels of genetic exchange are often referred to as **metapopulations**. It is important to note that the populations that are part of a metapopulation evolve together, and that changes which occur in some populations influence other populations. Evolution of multiple parts that are related to each other is called **coevolution**. Parasites and their hosts and predators and prey are other examples of coevolution. That does not mean that local adaptation does not occur for individual populations, but a population cannot be completely separated from the other populations to which it is connected; it and its genes influence and are influenced by other populations and their genes. Another way to think about it is that member populations become “acclimated” to the other populations, and to some degree the metapopulation itself evolves as an entity.

A recent example of how colonization occurs is the history of pink salmon colonization of the Great Lakes, which has many similarities to our model of postglacial colonization in Alaska (sidebar 3). The primary difference is that the colonization was the result of a single transplant, whereas in the North Pacific Ocean, there was opportunity for repeated contributions. As a result, it might be expected that the fish in the habitat colonized following the retreat of glaciers would be more similar genetically to the donor populations. Other than that, it appears that in the Great Lakes a relatively small number of translocated fish survived to reproduce. In addition, several generations elapsed before even modest numbers (relative to population sizes in their natural range) were reported, even in the streams near the release site in Lake Superior. The spread throughout the Great Lakes took longer yet. Nevertheless, the spread was probably extremely rapid relative to postglacial colonization because the streams in the Great Lakes were well established and not ephemeral; an enormous amount of suitable habitat was already available. The other facet of the Great Lakes colonization that parallels our model is that the genetic compositions of Great Lakes populations were quite similar. This suggests that as of the 1980s a single “stock” colonized the region or that there was still substantial gene flow, or genetic divergence would have been detectable. The genetic similarities among populations observed in the 1980s in the
Sidebar 3
THE INADVERTENT PINK SALMON EXPERIMENT IN THE GREAT LAKES

Although planned introductions of pink salmon into the Great Lakes between 1870 and 1922 were unsuccessful, an inadvertent introduction in 1956 succeeded. An intensive program (737,000 fish) was intended for transplant to Goose Creek, a tributary of Hudson Bay. The fish, whose origin was the Lakelse River, British Columbia, were cultured at a hatchery at Thunder Bay, on Lake Superior (Figure 4). The intent of the transplant was to provide a sport and commercial fishery for the people of Hudson and James bays. A small number of fry (100-350) were lost into Lake Superior while they were being loaded aboard a seaplane for planting, and approximately 21,000 fry were inadvertently released into Current River, a tributary to Lake Superior. The transplant to Goose Creek was unsuccessful; however, in 1959, two maturing fish were caught by anglers in two Minnesota tributaries. Only a handful of pink salmon sightings were reported until the fall of 1969 when pink salmon had spread to many Lake Superior tributaries. The genetic compositions of pink salmon populations derived from that release were examined more than 20 years later (1980-1985). The populations included in the study are shown on the map as circles.

Figure 4. The Lakelse River was the source of pink salmon inadvertently released into the Great Lakes at Thunder Bay in 1956. The genetic compositions of pink salmon populations derived from that release were examined more than 20 years later (1980-1985). The populations included in the study are shown on the map as circles.
What Does Genetics Have to Do with It?

In their natural range in the northern Pacific Ocean and Bering Sea, pink salmon have a rigid two-year, anadromous life cycle to which exceptions are exceedingly rare. The temporal separation that results from that life cycle has produced two genetically distinct broodlines, each spawning either in even or in odd years. In contrast, in the Great Lakes both even- and odd-year populations emerged from the single 1956 transplant. Three-year-old pink salmon are not uncommon in the Great Lakes and mixtures of mature two-, three-, and possibly one-year-old fish have been reported in some spawning populations. The genetic compositions (allozymes) of collections of pink salmon sampled from streams in Lake Superior, Lake Huron, and Lake Michigan in 1980, 1981, and 1985 indicate that they descended from the single lineage transplanted from the Lakelse River in British Columbia (Figure 4). Notably, the even- and odd-year samples do not differ, although even- and odd-broodline pink salmon from the same stream in their natural range differ substantially. In addition, comparison of allele frequencies from these Great Lakes collections with frequencies in the Lakelse River donor, show substantial differences. The nature of the differences suggests a strong founder effect (see Chapter 2. How Genes Vary in Fish Populations). At several allozyme loci, alleles that were rare in the Lakelse population were abundant in all of the Great Lakes collections. Those alleles are not abundant in the natural range of pink salmon. In addition, many other allele frequencies differed between the Lakelse and Great Lakes fish. It is apparent that only a small number of individuals, far fewer than the 21,000 or so fry that were released, were responsible for producing the Great Lakes lineage.

In the Great Lakes, there are now both even- and odd-year populations, which means that there must have been some three-year-old pink salmon. Both even- and odd-year populations have similar allozyme compositions, which means that they descended from the same inadvertent introduction, and there is no record of any other successful introductions. The appearance of three-year-old pink salmon fish in the Great Lakes, a breakdown in life history pattern, suggests that the Great Lakes do not have as rich an environment as the North Pacific Ocean and that some pink salmon are incapable of growing to a minimum condition to spawn in a two year period and do not mature until they are three years old. If this is true, it is possible that during and just following glacial periods or near the northern limit of their natural range, three-year-old pink salmon may be more prevalent.

tributaries and spawning runs as large as 1,000 had been observed in some Ontario tributaries. By 1973, they had been reported in nearly every suitable Lake Superior tributary as well as in Lake Michigan and Lake Huron. They were first noticed in lakes Erie and Ontario in 1979.

The rapid spread of pink salmon throughout the Great Lakes is not typical of salmon populations. This atypically rapid colonization may reflect either an increase in the proportion of fish that cannot find their way “home” to spawn, or alternatively a change in the cues such as currents and magnetic field that are involved in navigation processes, which are intrinsic to homing.

The Great Lakes lie outside the natural geographical and environmental range of pink salmon. The Great Lakes pink salmon are the only known self-perpetuating freshwater populations, which is notable because many attempts either to transplant pink salmon into lakes or to raise pink salmon of marine parentage in fresh water, without making dietary adjustments to compensate for the lower salinity, have had little success. Chinook, coho, and sockeye salmon as well as steelhead have all been introduced successfully to the Great Lakes. Efforts to introduce chum salmon failed, however, probably because the emigrant fry require a saltwater environment.
Great Lakes suggests that the influences of local adaptation in these relatively stable systems had not yet resulted in decreasing straying rates and producing genetic divergences. We suggest that these colonized populations are similar to what would happen following post-glacial colonization.

Summary
As recently as 14,000 to 20,000 years ago, glaciers covered much of the coast of the Gulf of Alaska, and ice extended out along the Aleutian chain. Over the past million years, similar advances of glaciers occurred about every 100,000 years. In fact, warm interglacial periods (like we enjoy now) have been rare. At the extremes of glaciation, species like Pacific salmon were displaced in many regions. In addition, the glaciers tied up much of the earth’s water and sea level was as much as 120-130 meters lower; this exposed the Bering land bridge, which linked North America and Asia together. Both the lowered temperatures and altered sea level drastically changed oceanographic conditions. Following the retreat of the glaciers, newly exposed streams were colonized by salmon and other species. The genetic compositions of many Alaska populations of salmon and other species still reflect post-glacial colonization patterns.
In this chapter, we examine the nature of several of the more commonly used kinds of genetic variation, and how that variation is detected in the laboratory. This chapter is important because it describes several of the methods that generate genetic data for Mendelian and population genetics analyses (Chapter 1. Even Fish Obey Mendel’s Laws, and Chapter 2. How Genes Vary in Fish Populations). One of the concepts that is taken for granted by population geneticists, but is not immediately obvious to others, is that the traits we score are merely tools used to learn about the variation in a population. They serve as indicators of the processes that have been going on, but their actual function is usually immaterial. In addition, from the perspective of the discipline, the gene pool (population) is the focus. We generally do not care about individuals, except as vehicles for the alleles in the gene pool. We usually look for loci that are representative of all the loci. Then we use their alleles as indicators to provide a window to the patterns of variation within and among populations or markers that, in some combination, reflect the results of the processes or track populations. For most basic questions, we are content with a set of anonymous loci; and in our initial surveys of a species, we treat each locus as anonymous, even if we know its function. Only in subsequent and more sophisticated studies might we ask questions about the role of particular loci. Those special loci often draw our attention because the way they are distributed among populations differs from the distributions of other loci and suggests that they may teach us something that the other loci cannot.
In chapters 2 (How Genes Vary in Fish Populations) and 3 (History of a Salmon Population), we learned that complex quantitative (= polygenic or metric) traits such as saltwater or temperature tolerance or timing are often the keys to local adaptation. We also learned that quantifying those traits generally requires breeding experiments, which cannot be easily conducted on wild populations. Breeding experiments are also usually time consuming and expensive. In addition, heritability estimates tell us that there is variability, but they do not tell us what the actual genetic (allelic) composition is. The short of it is that quantitative genetic traits are important, but they are impractical to apply in population surveys and may not provide the kind of data necessary for studying populations. For those reasons, most population geneticists focus on single (Mendelian) loci, which they use as markers (see Chapter 2) to provide a window to the genetic variation of populations. For such markers, even though the variants for many of them are neutral or nearly neutral (neutral means that natural selection does not recognize different phenotypes), the influences of historic relationships and of gene flow patterns and isolation (random drift) can be detected in the divergence patterns among populations. Variation at most of those loci is detected with biochemical methods. Recall that a locus and its alleles refer to a particular DNA sequence.

The entire set of genetic information, all of the DNA sequences of an organism, is referred to as its genome. All genetic variation ultimately comes from changes in DNA sequences. Most of us are aware that the human genome has been sequenced, as have genomes of numerous species that have been chosen as models for the vast diversity of organisms on earth. Much of the early population genetics work involved studies of variation in specific enzymes that are involved in metabolic processes. More recently, the DNA sequences themselves have been the focus of population geneticists to provide genetic markers. One of the objectives of the application of molecular genetics data to fisheries issues is that they may serve in some combination to delineate between populations, like the tattoos that are used to mark thoroughbred horses.

The DNA sequences that specify proteins and the varieties of RNA that are important in protein synthesis (as well as other types of RNA that have been discovered) are referred to as structural genes. There are many other genes, the most important of which are the genes that may not be expressed as structural products, but determine which structural genes need to be expressed at a particular time and which genes do not need to be expressed. These are regulatory genes. A rich example of regulation is the complex sequence of events that occurs during development of an organism from a fertilized egg through its emergence as a fully formed individual. This development process is choreographed by regulatory genes.

In addition to the genes that have obvious roles, there are many genes that have no apparent role. Examples are the microsatellite loci, which are described in this section, and other short sequences that may be scattered throughout the genome such as the Alu sequences in humans. The Alu sequences are about 200 nucleotide pairs long (the entire genome is about 3 billion nucleotide pairs), and there are about 1 million copies of Alu sequences distributed throughout our genomes. That is, about 10% of our DNA sequences are this short sequence repeated over and over. Most other species do not have these Alu sequences, but many have large numbers of other small sequences. These kinds of sequences have earned the name “junk DNA.” But we need to be careful. Just because we do not know of a function, does not mean they do not have one. For example, it was discovered recently that a few microsatellite loci play roles in regulation. Finally, in some instances, the term locus may refer to a single nucleotide site and the different nucleotides that occur at that site are distinct alleles.

Before we examine some of the methods that can be used to gather genetic data, it is important to stress one particular point. Application of genetic information to fish stock identification questions requires baseline information from most (preferably all) of the populations that may contribute to a mixture. I cannot begin to tally all of the instances during my career that an individual or agency has wanted to provide me samples, which they had taken from a catch of one or another species, so that I could tell them where the fish came from. In nearly every instance, they were disappointed when I told them
that I couldn’t do it because there was no baseline information available to serve as references for the genetic compositions of the potential contributors to the mixture for that species. They were even more disappointed when I told them that assembling a baseline requires substantial numbers of samples from spawning populations. The costs of mounting the field operations and then analyzing samples in the lab are not trivial. As a result, most genetic studies involve species that are important commercially or that have severe conservation problems. So, what can we do? Read this and the other chapters for answers!

Most biochemical genetic markers reveal variation in the DNA sequence, either directly or indirectly. Although DNA sequencing provides the most detailed data on variation, it is still impractical to sequence the number of loci and individuals that would be necessary to characterize the population structure of a species, in order to discover markers that provide adequate resolution for stock identification applications. In this chapter, we discuss four methods: allozyme analysis, restriction fragment length polymorphism (RFLP) analysis, microsatellite analysis, and single nucleotide polymorphism (SNP) analysis. We describe the nature of the variation that they resolve and the way in which the variation is detected. Of course, there are other methods that resolve genetic variation, but the ones considered here are the most widely applied.

**Allozymes**

Many genes carry information for proteins. Proteins are large molecules that have many different functions in an organism (see sidebar). Those functions include structural (e.g., muscle components) and communication (e.g., hormones like insulin or growth hormone). Enzymes, one of the most important kind of protein, mediate (catalyze) all of the biochemical reactions that take place in an organism. Enzymes are responsible for metabolism. Proteins are long chains of amino acids. Amino acids differ from each other because they carry different small chemical groups. The sequence of nucleotides in the DNA directs the order of assembly of amino acids into protein. A single change in the DNA sequence can cause the replacement of one of the amino acids by another.

Two alleles at a locus can carry subtly different versions of the instructions and can specify slightly different amino acid chains (Figure 1 top). Dissolved proteins have an intrinsic electrical charge that is related chemically to the amino acid composition. Some mutations can produce amino acid substitutions that create small electrical charge differences between their protein products. The differently charged proteins can be physically separated in a direct current electric field by a technique termed protein electrophoresis. Allozymes are electrophoretically detectable gene
products of alleles at a locus; different allozyme alleles have the same function but different mobilities. Ordinarily, allozyme loci produce only a few detectable alleles.

Allozymes were the mainstay of population genetics for nearly four decades, but few labs continue to conduct allozyme studies. Allozyme analysis indirectly detects variation in the DNA sequence by looking for differences in the proteins that the gene specifies. Many of the proteins that are analyzed are enzymes. Allozyme analysis focuses on proteins (mostly enzymes) that are dissolved in the cytoplasm of a cell.

Protein electrophoresis provided the first tool that was practical for generating large amounts of data that could be used to study the population genetics of a wide variety of species. It was the dominant method used from the early 1970s until the late 1980s. It was applied generally to population genetics questions, to determine parents in limited experiments, to detect hybrids, and in stock identification. At the time, protein electrophoresis had the advantage that it was inexpensive, required little specialized (i.e., expensive) equipment, and could rapidly be used to obtain data from multiple loci for many individuals. Unfortunately, much of the data required heart and liver samples, which depended on lethal sampling. Because many of the enzymes analyzed are sensitive and easily lose their activities, the tissues need to be stored at very cold temperatures (−80°C is typical) soon after they are taken, which is impractical in many remote areas. In addition, only a very specialized portion of the genome is examined (central metabolic pathway enzymes), and protein electrophoresis resolves only the mutations that change the net charge of the protein. Other methods that have been developed more recently resolve more variation than allozymes. In addition, the cost of protein electrophoresis has increased substantially in recent years, and not very many labs still conduct allozyme analysis.

During the past decade or so, methods that were developed to detect allelic differences in DNA nucleotide sequences have been applied extensively. The other three methods that we use to detect genetic variability involve DNA analysis. The innovation that has made these methods practical is the polymerase chain reaction (PCR). Another advantage of DNA-based methods is that samples of tissue (a clip of a fin or even scales) can be taken for genetic analysis without killing the organism.

**The polymerase chain reaction**

The information carried in the DNA sequences of vertebrates, such as humans and fish, is encoded by...
about 3 billion pairs of DNA nucleotides. The loci that are analyzed to provide population genetics data are a tiny portion of that total. In fact, a major constraint to studying them is the “needle in the haystack” problem. Target genes make up such a small portion of the total DNA that they can be difficult to detect. So, the first step is to generate enough of the target DNA sequence to study. Early in the development of DNA methods, DNA or gene cloning was used to generate the material necessary. Although cloning is still an important tool for some molecular genetics studies, it is too tedious and time consuming to be practical for population genetics surveys and has been supplanted by the polymerase chain reaction as a method to provide huge amounts of a particular DNA sequence for study.

The polymerase chain reaction is a way to make enormous quantities of (amplify) a DNA sequence that we want to study, the target sequence. It involves synthesizing DNA in a test tube, but the trick is that only the target sequence is copied. DNA is separated by heating (melting) into its two complementary strands (see sidebar 1). Small sequences (called primers) that are complementary to sequences flanking the target sequence stick to them (anneal) and provide an opportunity for a DNA synthesis enzyme (polymerase) to make a new complementary strand. The target sequence is replicated (but not the rest of the genome), so now there are twice as many of the double strands of the target sequence. The new DNA is again separated into its complementary strands, the primers anneal to the single strands, and DNA is synthesized again to double the number of target sequences. There are now a total of four relative to nontarget sequences. The process is repeated and the target sequences are doubled each cycle.

The polymerase chain reaction (PCR) uses components that can be purchased commercially. Remember that we are focusing on a specific target sequence, and a major drawback is how to select that region specifically. It turns out that this can be a challenge. What we need is information about the DNA sequence that flanks the target sequence. We usually get that information from sequences of DNA cloned from the species that we are studying. Alternatively, we look at DNA sequence data banks, which often have DNA sequences of the target gene for other species. Sequences that are very similar among species (conserved sequences) might be expected to be similar in the sequence of our species (but there are no guarantees). Once we have that information, we can design short (~20 nucleotides) single-stranded DNA fragments, called primers, that complement the DNA sequences flanking the target region. It turns out that when we heat DNA, the two complementary strands separate. If we have primers, we can make new strands to complement the exposed single strands.

The PCR process separates the double-stranded DNA into its two complementary strands when the DNA is heated to about 95°C. The temperature is reduced; and short nucleotide sequences, which are designed specifically for the sequences that flank the target region, pair with their complements on the exposed single strands and “prime” DNA synthesis. Completion of the DNA synthesis doubles the number of copies of the target sequence. We now repeat the process. The temperature is again elevated to separate the double strands of DNA. These steps are repeated and the quantity of target sequences doubles at every cycle. After thirty cycles, this PCR amplification yields $2^{30}$ copies, more than 1 billion times more (Figure 3). What makes this
process possible is a commercially available heat-stable enzyme that synthesizes DNA. Most proteins denature at elevated temperatures (greater than about 60°C). Without this heat-stable enzyme, Taq polymerase, which was isolated and cloned from a hot spring microbe, fresh polymerase would have to be added each PCR cycle.

**Restriction site analysis**

Restriction endonucleases—also referred to as restriction enzymes—recognize specific short (often four to six nucleotides) sequences of double-stranded DNA. They cut those sites surgically like a pair of molecular scissors. Where do they come from? They come from bacteria. Restriction enzymes are used by many bacteria to destroy “foreign” DNA, such as the DNA from viruses that may try to infect them. These enzymes are called restriction enzymes because they “restrict” the DNA that is allowed to exist in the bacterium. How do the bacteria protect themselves from their own restriction enzymes? They have a second enzyme that chemically modifies a nucleotide in the restriction site to “immunize” it from the restriction enzyme. In newly synthesized DNA, one of the molecule strands has modified sites, but the other strand doesn’t. The second enzyme recognizes the modified site on the parental strand, and uses it as a guide to determine where to modify the newly synthesized strand. These restriction enzymes surgically cleave DNA and are used to detect those specific target sites in DNA samples. The result is a set of precisely sized fragments. Molecular biologists have “borrowed” restriction enzymes and “harnessed” them for laboratory use (they can be purchased commercially). This battery of restriction enzymes can recognize many different nucleotide sequences. Many allelic variants in the DNA sequence can be detected indirectly by restriction fragment length polymorphism (RFLP) analysis. Target sequences (often the result of PCR amplification) are digested with restriction enzymes. Restriction endonuclease digestion repeatedly produces a set of fragments whose lengths depend on the nucleotide sequence. An alteration (mutation) in one nucleotide of an endonuclease recognition site will prevent its cleavage and alter the restriction fragment set (Figure 4). After endonuclease digestion, the restriction fragments can be electrophoretically separated by size in agarose or polyacrylamide gels. Like the starch gels for protein electrophoresis, these gels are similar to gelatin desserts. They provide support for the DNA fragments, but permit movement of the fragments toward the electrode that has a charge opposite those of the DNA fragments. The positions of the DNA
fragments in the gel can be detected with DNA-specific dyes (Figure 5). The restriction fragment data are not directly applied to most questions, but they can be used to map the location in the DNA sequence of the restriction sites that produce them. Analyses are usually conducted by using the presence or absence of the restriction sites.

All three of the DNA methods have the advantage that only small samples are required and that an individual can be sampled nonlethally. In addition, the tissue sample can be preserved in several different ways that do not require freezing. DNA methods can often take advantage of samples that have been archived from other studies, such as allozyme samples or scale collections. The DNA methods are also based on PCR amplification. The advantage of RFLP analysis is that the restriction endonucleases can sample DNA in sequences that are too large to be practical to sequence. Restriction site analyses have often been applied to mitochondrial DNA (mtDNA). Although mtDNA data have been applied to many kinds of studies, two of the more important applications have involved phylogenetic questions and the recent (in a geologic time frame) demographic history of a species. Some of the more interesting studies address questions of post-glacial colonization.

The disadvantages of RFLP studies are that, compared to sequencing, the approach sacrifices the “fine detail” that is in the complete nucleotide sequence of a relatively small region of DNA for the less detailed information of a much larger DNA sequence or many more individuals. It also requires more specialized equipment than allozyme analysis. Finally, mtDNA is a single intact molecule, essentially a single locus; it is inherited clonally from mother to offspring, so it is haploid (like a gamete). Most stock identification methods depend on diploid multi-locus data for fine resolutions.

Microsatellites

Alleles of microsatellite loci are another type of DNA marker that is frequently used. Microsatellites are back-to-back repeats of short (about two to four) nucleotide sequences. One of the common repeats is CACA. The number of repeats can range from a few to more than a hundred. We define microsatellite alleles by the number of repeats or size measured in nucleotide pairs. Size differences are detected by the same technique as for RFLP—by electrophoresis.

Microsatellites are detected by using PCR to amplify the entire set of repeated sequences that lies between the nonrepeated flanking areas. The sizes (number of nucleotides) of the microsatellite alleles are estimated from their electrophoretic mobility, as described for RFLP analyses. Many PCR microsatellite alleles are relatively small (100 to 400 base pairs) and are resolved by an automatic DNA sequencer (see an example of a microsatellite gel in Figure 6). Although they may seem a bit unusual, microsatellites are abundant in most animals and occur along chromosomes about every 10,000 nucleotides or so. Of course, each microsatellite locus is flanked by different sequences and may be made up of different repeated sequences. Not only are there large numbers of these loci, but a single locus may have as many as 100 different alleles. In contrast, allozyme loci generally have only a small number of alleles (two is the most common, but several are observed at some loci), and single nucleotide polymorphism sites (see below) generally have just two alleles. In addition, the mechanisms by which new alleles are produced occur more frequently (often 100 times or more as often) than mutations at most other loci. Most microsatellite

![Image of microsatellite variation in shortraker rockfish (Sebastes borealis). Each fish occupies a single vertical lane, except lanes 1, 18, 35, and 52, which have size standards. Most of the fish are heterozygotes (two bands) stacked one over the other.](image-url)
DNA, RNA, AND PROTEIN—A GLIMPSE OF MOLECULAR BIOLOGY

DNA (deoxyribonucleic acid) is the molecule that encodes information needed for every facet of an organism’s existence. DNA includes directions for structures, functions, and even behaviors. DNA is a linear molecule, which is made of subunits called nucleotides attached together like a very long train. A nucleotide (formally a deoxyribonucleotide) is built from three pieces: a phosphate, a simple five-carbon sugar (deoxyribose), and a flat organic ring structure of nitrogen and carbon referred to as a nitrogenous base or base (Figure 7).

There are four different bases in DNA: adenine (A), thymine (T), cytosine (C), and guanine (G). Figure 8 shows an example of a guanine nucleotide (formally guanosine monophosphate).

The phosphates and deoxyribose sugars alternate in a long DNA chain, and the bases stick out from the deoxyribose sugars (Figure 9).

The As and Ts have an affinity and can bind weakly to each other. Similarly, the Gs and Cs also attract each other. The result is that DNA is a double-stranded molecule, a duplex (Figures 10 and 11).

In their original description of the structure of DNA, Watson and Crick slyly pointed out that “it had not escaped their notice” that each strand on its own can direct the synthesis of its complementary strand when DNA is replicated (the process by which one parental DNA duplex produces two identical daughter DNA duplexes). Replication precedes mitosis and meiosis (see Chapter 1, Even Fish Obey Mendel’s Laws).

DNA duplexes wrap around each other to form a helix. The phosphate-deoxyribose chains are represented by the two gray ribbons on the outside of the structure in Figure 12.

DNA molecules in organisms are much longer than the figures shown above. How long are they? The bacterium Escherichia coli, E. coli for short, has about 3 million nucleotide pairs. We humans have about 3 billion nucleotide pairs. The total length of DNA in a human is about 6 feet—in each cell! That is one of the reasons that the DNA is broken into shorter pieces (chromosomes) for cell division. The length of the DNA molecule in each of 46 human chromosomes is about 1½ inches, still much longer than a cell (about 0.0004 inches).
Many genes are encoded in the DNA of organisms. In addition to structural genes, there are many sequences that have no known function—maybe half or more of the total DNA! The primary products of genes that carry information in the DNA sequence that directs construction are a diverse set of proteins. Proteins include an array of functions. Structural proteins are parts of muscle, bone, and cartilage. Catalytic proteins (enzymes) mediate the biochemical reactions needed for synthetic metabolism (such as amino acid synthesis), and for energy-deriving metabolism (such as converting fats and sugars to energy).

DNA is packed into the nuclei of organisms that have them. However, proteins are constructed outside of the nucleus in the cytoplasm, where protein synthesis is conducted by structures called ribosomes. The DNA serves as the central data bank for information. In order to construct a particular protein, a copy of the instructions for that protein is made, and those instructions are transported to the ribosome in the cytoplasm. The instructions are carried by ribonucleic acid (RNA) molecules, which are very similar to DNA. RNA has a ribose sugar instead of a deoxyribose, uracil (U) instead of thymine (T) bases, and is (usually) single stranded. There are several functionally different RNAs; the ones that carry instructions for constructing proteins are messenger RNAs (mRNA). DNA directs the synthesis of RNA, which is complementary to the DNA strand from which it was made; uracils (U) and adenines (A) are attracted to each other, so Us are replaced by Ts in the RNA copies (Figure 13).
Two other kinds of RNA are also involved in protein synthesis. Ribosomes are made from proteins and special RNAs called ribosomal RNAs (rRNAs), which actually participate in the protein synthesis. Another essential RNA, which interprets the information carried by the mRNA sequence, is called a transfer RNA (tRNA). The code carried by the DNA and passed to the mRNA is a language of three-letter words. Three consecutive nucleotides (called a codon) spell a word for a particular amino acid. Note that the tRNAs also use the complementarity between bases to mediate the correct placement of amino acids in the new polypeptide.

The linear DNA and mRNA codes direct the construction of a specific linear order of amino acids, or polypeptide. Each amino acid has a unique chemical nature. The function of a protein is determined by its amino acids. Below we can see that the codon GCG on the mRNA specifies the amino acid alanine. Similarly, the amino acid glutamine was specified by CAG; leucine was specified by CUU; and the next amino acid that will be incorporated, serine, is specified by AGC (Figure 14).

loci are not part of genes that are expressed as proteins; however, some microsatellites may be involved in determining if and when other genes get expressed.

Microsatellites are advantageous for general population genetics studies, which include development into stock identification tools. Microsatellite variation is also used to determine who the parents are in a variety of situations, such as breeding studies in which offspring are released to the sea and recovered when they return. Microsatellites have also been used in genome mapping projects. Many microsatellite loci are available and the primers that are designed for one species often amplify loci in related species. Many loci have multiple alleles, which can be advantageous for stock identification and parentage studies.

There are also several disadvantages associated with microsatellite analysis. The equipment can be expensive; for example, DNA sequencers are often used to get the data. If primers are not already available, development of primers can require some expertise and effort. Clear, readable gel patterns are not always easy to obtain and some loci may require effort to standardize between labs. Finally, mutations in the primer sequences can cause the PCR reactions to fail. These failures are referred to as “null alleles.” Unfortunately, these failures cannot be easily detected and can bias estimates of allele frequencies.

**Single nucleotide polymorphisms (SNPs)**

Microsatellites sample variation throughout the nuclear genome, but from very distinct kinds of genes. It would be nice to be able to sample variation more broadly throughout the genome without having to sequence the DNA. When you get down to it, much of the variation that we would like to detect for population genetics studies results from single nucleotide changes in the DNA sequence (Figure 15). For example, the difference between two allozyme alleles usually results from a single nucleotide change, as does much of the other variation, even
variation that occurs in genes other than protein-coding genes. These simple DNA sequence differences are **single nucleotide polymorphisms** or **SNPs**. Clearly, if we could detect variation at the level of the SNP, we would probably have all of the variation that we need to address many problems. The data that would be useful for population genetics, or stock identification, would be taken separately for each locus (or SNP site) for each individual. Those data could then be analyzed by using the same tools that are used for allozyme or microsatellite data.

To apply SNP analyses to population genetics questions requires several steps. First, to address a specific question in a previously unstudied species, it is necessary to “discover” SNP sites in the DNA sequence. There are several ways to do this. The discovery process can involve searches by sequencing genes in a set of organisms that we wish to study. Alternatively, we can mine DNA sequence data banks like we did to develop primers for PCR amplification. The next step is to design a method (an assay) that detects a target SNP site and differentiates between each distinct sequence. Several different methods are available that can “genotype” SNP sites once a specific assay has been designed for the particular detection method, and more methods will surely be developed. A currently popular method is the Taqman™ assay, which takes advantage of the imperfect matching of synthetic, short, single-stranded nucleotide sequences, called “probes” that are designed to detect a specific SNP site. It also uses an enzyme for DNA synthesis that has a DNA polymerase activity at one end and a DNA digesting activity at the other end. The probes are added to a PCR reaction that amplifies a short sequence, which includes the SNP site. Probes that are not exactly complementary to the SNP site are displaced by the polymerase; a probe that is a perfect match is digested by the enzyme because it is in the way. The digestion releases a fluorescent dye that is detected by the instrument. The amount of dye in a reaction increases with each iteration of the PCR reaction (Figure 16) until there is enough to be detected. Individuals are analyzed one by one; an individual that is homozygous will produce only one color, but a heterozygote will produce both colors. An instrument is used to detect the presence of dyes and report the results. Frequently small trays

---

**Figure 15.** A single nucleotide polymorphism (SNP) is shown in fish no. 2, where a “t” occurs instead of a “c” in the DNA sequence.

**Figure 16.** The Taqman™ assay uses two “probes” that have different fluorescent dyes. One homozygous single nucleotide polymorphism (SNP) produces a signal from one of the dyes, but a heterozygote produces signals from both dyes.
of 96 (12 × 9) or 384 (24 × 18) individuals are scored by the instrument, and the genotype of each is recorded electronically. The Taqman™ assay is only one of several detection methods. One of the great advantages of SNP variation is portability among laboratories and the potential to adapt the variation to alternative detection methods.

Much of the genome can be sampled for SNP variation. The variation is simple to interpret because there are usually only two variants at an SNP site. In addition, the processing assays can be readily adapted to automation, including scoring the variation. The disadvantages are that SNPs must be newly discovered for each species to which they are applied and specific assays must be developed for each new SNP. SNPs generally have only two “alleles” per site, which is much less information than carried by the multiallelic microsatellites. Ordinarily, many SNPs would need to be applied to questions that require a dozen or so microsatellite loci. Finally, if more than one SNP is detected in the same region of a gene, they cannot all be used directly as multiple SNPs. Nucleotides that are physically linked tend to be inherited as a unit, not independently. Linkage can cause problems with interpreting the data. Finally, SNPs are a relatively new method. Start-up is expensive and baseline information, which must be developed if the genetic information is going to be used for stock identification, must be gathered before the tool can be applied.

Regardless of its current disadvantages, SNPs or a method based on SNPs will probably supplant most of the other currently used genetics methods in the not-too-distant future. One of the possibilities is a “DNA chip”—called a micro array—that can genotype many loci simultaneously. The challenges are that (1) a different chip would probably be required for each species, (2) for most species there are not yet a sufficient number of loci that provide discrimination among populations, (3) the chips are expensive, and (4) for now the chips can be used just once.

Summary

Four methods for detecting genetic variation were presented: allozyme analysis (protein electrophoresis), restriction fragment length polymorphism (RFLP) analysis, microsatellite analysis, and single nucleotide polymorphism (SNP) analysis. Allozyme analysis indirectly detects DNA sequence variation from the changes in electrophoretic mobility that result from changes in amino acid compositions. Allozymes are the products of structural genes. The other three methods directly analyze DNA variation and usually rely on polymerase chain reaction (PCR) amplification to provide large quantities of target DNA sequences for analysis. Restriction fragment analysis is based on changes in the recognition sequences of restriction enzymes and can be used to sample large sequences; it does not, however, provide complete sequence information. RFLP variation can occur in any DNA sequence, structural genes, or noncoding genes. Microsatellite analysis focuses on an abundant type of locus that is spread widely through the genome. These loci have large numbers of repeats of very short (two to four) nucleotide sequences and their alleles vary in the number of repeats. Microsatellite loci are generally noncoding genes, but some may be involved in regulation. Finally, SNP sites are simply single changes in the nucleotide sequence. A SNP can be anyplace in the genome.
In previous installments we introduced some of the concepts and jargon that are needed to navigate through genetics (Chapters 1. Even Fish Obey Mendel’s Laws, and 2. How Genes Vary in Fish Populations), we speculated about the origins of present-day salmon populations in Alaska (3. History of a Salmon Population), and we described some of the molecular and quantitative methods that are used to obtain data for studying the genetics of populations (4. Molecular Tools for Population Genetics). The objective of this chapter is to look more closely at the processes that influence the genetic structure of existing salmon populations, and to examine in detail the genetics of salmon in a small stream that we have studied intensively—pink salmon in Auke Creek near Juneau, Alaska.
Recall that a population (the way we have been using the term) is a spawning aggregation. In an idealized population, each male has an equal chance to mate with any female. Yes, I know that in reality the mating schemes are much more complicated, but it is a good place to start from and probably not too far off.

The first thing we need to realize is that populations and their genetic compositions are dynamic, so we are tracking a moving target. But with that said, the dynamic feature is part of what we need to understand. As has been the case for most of their histories, salmon populations encounter environmental changes daily, seasonally, yearly, in cycles of several years that influence the climate, and in cycles that are measured in geologic time that produce advances and retreats of glaciers. In addition, no two years bring exactly the same conditions throughout the geographic or oceanographic range of a particular population. From a natural selection perspective, that means that allele frequencies get jerked around a bit every generation.

Most organisms pass through a series of transitions between fertilization and death. Some of the stages in salmon are developing embryos in the gravel (or hatchery incubator), emergence from the gravel, freshwater life (most pink salmon have none), emigration to salt water, early growth in the estuaries, migration to the open ocean, maturation and return to natal streams, spawning, and ultimately death. These stages are referred to as its life history. Most of the traits that are important to a fish’s life history are quantitative traits, the combined result of effects of alleles at many loci. That means that for the most part alleles at a single locus are not being singled out for strong selection. It also means that much of the genetic variation for a trait remains after directional selection over longer times, one would ultimately expect large changes in both phenotype and genetic composition. The year-to-year changes in environment have a much stronger random element. Some years may nudge the phenotypic (and underlying genotypic and allelic) distributions in one direction, and other years might nudge them in the opposite direction. As a result of this back and forth movement, alleles that produce extreme phenotypes will tend to be weeded out but
variability will be maintained in large populations. It is important to keep in mind that a broad cross section of the population will contribute to each generation, but they contribute to different degrees and the phenotypes that are most successful one year will probably not be as successful in other years.

The variability that exists in a population will reflect its recent environmental history. What is important is that the genetic variability that exists in the population usually produces some phenotypes that are not at the center of the distribution. Those phenotypes may allow for at least some members of the population to survive some extreme environmental fluctuations. In the long term, variation within and among populations will provide the evolutionary fodder for inevitable global changes. Species, or populations, that have little genetic variation will likely join the rapidly growing list of those that have gone extinct, and that event will probably occur sooner rather than later.

Just as populations adapt to changes on a geologic time frame, they also adapt to changes on a manager’s time frame. Some of those changes are simply physiological responses to temperature and other environmental conditions. Most fish can tolerate a range of temperatures, salinities, food availability, and even toxicants. Such internal physiological adjustments are often referred to as plasticity. Because of built-in plasticity, some environmental fluctuation can be met by most individuals of a population. We just argued that even subtle environmental differences can alter allele frequencies (natural selection), but we also said that the amount of change in allele frequencies in one generation is probably small. We can also argue that the advantages of plasticity would make it a trait that is also under selection and that the amount of plasticity that exists in individuals in a population may to some extent reflect the population’s history.

Most salmon populations are not completely separated from other populations. We described the dynamics between populations previously (Chapter 3. History of a Salmon Population). The concept we presented there was metapopulation. Part of the idea was that many populations within a geographic area were probably seeded originally by colonists that shared the same gene pool. The rest of the idea is that genetic material is exchanged between populations by strays (migrants). We suggested that straying might be subject to natural selection and that populations that are in stable streams might have lower levels of straying than populations that occur in less stable streams. Regardless, for every population there is an advantage to having at least a low level of straying. This means that within a geographical area, there is probably limited local exchange of alleles. The frequencies of alleles would respond to the selection pressures within a stream, but they would also respond indirectly and to a reduced degree to the environments in other populations connected in the metapopulation. However, an allele that is severely disadvantageous in a particular stream will be quickly eliminated in that stream, whereas others will persist longer. What this means is that crucial variability within populations can be maintained generally by the exchange among populations at the same time that maladaptive alleles are eliminated. For an allele that is maladaptive in one population to be able persist in the metapopulation, it must have an adaptive advantage in another population. Recall also that in geologic time, individual populations occasionally disappear as a result of catastrophic events. But if there is useable habitat, it will often be reseeded from other local populations; and generally the genes it is seeded with have a reasonable likelihood for success because most represent the general local area, including the stream from which they disappeared. Of course, restoration may require a long time (in a manager’s time frame) to occur.

By now you should have the general flavor of the relationships among populations as well as the nature of genetic responses of a population to selection. A single stream may have several populations (in our narrow genetics definition). For example, salmon may return to different locations or habitats within the stream or they may return at different times. I will describe some of the research that we (my colleague Bill Smoker and many graduate students) have done on pink salmon in the Juneau area, especially Auke Creek, that is pertinent to population structure. We started our work on Auke Creek pink salmon in the late 1970s with a genetic marking experiment. Genetic marking is accomplished by changing the allele frequencies at a locus in one population relative
What Does Genetics Have to Do with It?

to the frequencies in another. By breeding only fish that carried a specific allele (which we will refer to as a marker allele), we changed the frequency of an allozyme (see Chapter 4. Molecular Tools for Population Genetics) allele in one of the spawning aggregations in Auke Creek. Many of the alleles of biochemical genetic traits, like the allozymes we used then, are neutral or nearly neutral. That is, natural selection has little influence on those traits. As a result, fish that stray from one population to another (see Chapter 3. History of a Salmon Population) tend to homogenize populations (give both of them the same allele frequencies) for these loci. Recall, however, that at the same time populations are being homogenized for allele frequencies at neutral loci, local environmental differences can cause divergence at other loci that are important in local adaptations, and random drift can make random changes in allele frequencies.

Auke Creek is a short stream that flows about 400 meters from Auke Lake to Auke Bay. The National Marine Fisheries Service Auke Bay Laboratory (ABL) operates a weir (a fish fence) and small research hatchery just above tide water (Figure 1). The weir can be configured to sieve returning adults in the summer and fall and emigrant fry (young fish) in the spring. Before we began our experiments, the Auke Creek Hatchery manager, Jerry Taylor, noticed two distinct runs of pink salmon; the first peaked in mid to late August and the second in September.

Figure 1. Pink salmon streams near Juneau, Alaska. The Auke Creek system is shown in the right inset. The location of Auke Creek is designated by 4, Lake Creek by 13, and Waydelich Creek by 3.
We genetically marked the late part of the late-run pink salmon in Auke Creek. After preliminary surveys of the allozyme compositions of Auke Creek and other streams in the Juneau area, we chose loci that code for the enzyme malate dehydrogenase, which we abbreviate as MDH (the formal name of the locus is $MDHB_{1,2}$). MDH is a central metabolic pathway enzyme involved in converting carbon compounds into an energy form that the cell can use. MDH is expressed in many different tissues. The particular allele we chose moved about 70% as fast as the most common allele at that locus in electrophoresis gels. We will refer to it as the *70* allele; the common allele is the *100* allele. Remember that we are using these as markers and as long as the alleles are relatively neutral, it does not matter which alleles a pink salmon has (although we should check out the neutrality at some point). Our surveys showed us that marker allele $MDH^{*70}$ had a frequency of about 5% in Auke Creek, and in the Juneau area in general; that is, about one in 20 alleles was the *70* allele. We spawned only the fish that carried the *70* allele. Because salmon are diploid, roughly 1 in 10 fish in the population carried the *70* allele. The process that we used for marking is shown in Figure 2. Most of the fish that carried the allele were heterozygous (*70/*100), depicted by the green fish above and the three-banded patterns in the gels, as compared to the common homozygous (*100/*100) fish, which are the blue fish and the single band closer to the top of the gel. Crossing the heterozygous breeders produced some fish that were homozygous for the marker allele *70/*70, which are represented by the yellow fish and the single bands closer to the bottom of the gel. Of course, the fish really did not have the color differences or we could have used the colors to mark the fish!

Also seems likely that such obvious differences would be fodder for natural selection. Think of that bird looking for a meal.) Among the fish that were screened for markers, the fish that carry the $MDH^{*70}$ allele are represented by the green fish above and the three banded patterns in the gels, as compared to the common homozygous (*100/*100) fish, which are the blue fish and the single band closer to the top of the gel. Crossing the heterozygous breeders produced some fish that were homozygous for the marker allele *70/*70, which are represented by the yellow fish and the single bands closer to the bottom of the gel. The genetic marking process increased the marker allele frequency from 0.056 in the stream to 0.508 in the marked lot of fish. When the marked fish returned as adults, they mixed and interbred with fish produced naturally in Auke Creek. The result was that after marking, the late run had a frequency of 0.253, substantially higher than the other spawning aggregations in Auke Creek and the other streams in the Juneau area.
What Does Genetics Have to Do with It?

There is usually a short time between the early and late runs during which no fish pass the weir; and many of the late-run fish have bright scales and are not yet able to tolerate fresh water, which indicates that they are recent arrivals. In addition, both runs appeared to segregate into intertidal (below the weir) and upstream (above the weir) aggregations. In some years, pink salmon also spawn in Lake Creek, a tributary to Auke Lake. The run timing for the Lake Creek spawners is not as distinct. The pink salmon stream nearest to the Auke Creek system is Waydelich Creek, about 1 km away.

In 1979, we marked the end of the late-run Auke Creek pink salmon that passed the weir from mid to late September. Our original objective was to see if it would be feasible to genetically mark hatchery stocks so that they could be distinguished from wild fish. If it were possible, we might be able to study hatchery-wild fish interactions. We did the marking by breeding only fish that had a particular allele at an allozyme locus that we could detect electrophoretically (see Chapter 4. Molecular Tools for Population Genetics, and sidebar 1). Although we were successful in marking the fish, the most important result of the experiment was that the genetic mark allowed us to identify and follow a particular spawning aggregation over several generations.

After we marked the late-run, upstream spawning aggregation, we sampled adults returning to Auke Creek for five generations (Figure 2). The frequency of the marker remained relatively (statistically) stable in both the marked late-upstream and unmarked early-upstream spawning aggregations during that period. Because very few fish returned to spawn in the intertidal area during this period, we stopped sampling there. The frequency of the marker in the late-intertidal fish increased somewhat, but did not reach the frequency of the late-run upstream fish. The marker frequency on the Lake Creek fish also increased somewhat. We monitored nearby Waydelich Creek, which had no increases in either the early or late runs. Note that all of the pink salmon production after 1979 was natural production from Auke Creek.

Our results showed that spawning aggregations within Auke Creek are separated both in time and space. The exchange of genes between early and late runs was very low (at least during this period of time); and there appeared to be only limited exchange between the spatially separated aggregations, even over as short a distance as 50 feet. We didn’t see any detectable exchange with nearby Waydelich Creek, either. What about Lake Creek? Well, the relationship between Lake Creek and mainstem Auke Creek is less clear. We did see an increase in the frequency of the marker in Lake Creek, but because the fish destined for Lake Creek had to pass the weir to get there, we may have included a few Lake Creek fish as breeders for the marker at the first generation. In some years there are virtually no spawners in Lake Creek; in other years—especially years of abundant returns or high stream temperatures—Lake Creek may accommodate overflow. In one especially warm year, fish moved rapidly through Auke Creek, held in Auke Lake until temperatures were more favorable, and then spawned in Lake Creek. We have too little data about Lake Creek to speculate any further.

We showed in Figure 2 that the marker is heritable and differentially marked the late upstream fish produced in our genetic marking experiment. There was no statistically detectable difference among the breeders, a sample of fry tested at release, or the returning fin-marked fish. We found no evidence of any alterations in fitness that might be indicated by changes in allele frequencies.

---

**Table:**

- **We fin-marked about 60,000 of the 178,219 fry that we released.**
- **The fin-marked fish provided us with a way that we could unequivocally identify genetically marked fish when they returned. Note that about ¼ of the lot that we released carried no *70* allele. Comparing the marker allele frequencies of the parents, released fry, and returning adults made it possible for us to determine if there was a detectable change in the frequency of the marker. A change in frequency between release and return in the marked fish could indicate selection against the marker allele (remember I said we would have to check this out). We recovered 1,048 fin-marked fish. These were unequivocally fish produced in our genetic marking experiment. There was no statistically detectable difference among the breeders, a sample of fry tested at release, or the returning fin-marked fish. We found no evidence of any alterations in fitness that might be indicated by changes in allele frequencies.
fish. Because it is genetic, the fish carried the mark throughout their lives; and we were able to use the marker to determine the timing of both adult returns and fry emigration by sampling 10 to 20 fish each day at the weir during the migration periods and looking for the marker. The return distribution was not surprising; the frequency of the marker began to increase in mid September at about the same time we had begun screening fish for marking. The surprise was that fry from early-run fish tended to leave Auke Creek before fry from the late-run fish (Figure 3).

The cumulative returns in Figure 3 also show that the number of early, unmarked fish exceeded the number of late, marked fish. It also indicates that the survival of late fry exceeded the survival of early fry, and that approximately equal numbers emigrated. Again, recall that all of the pink salmon production after 1979 was natural production from Auke Creek. We conducted similar experiments in three consecutive generations. It appears that during those years early-run fish had higher marine survivals, but late-run fish had higher freshwater survivals. We will return to this observation later.

Genetic marking revealed strong temporal structure in Auke Creek pink salmon, as well as spatial separation between streams. We also saw that there was possibly weak spatial structure within Auke Creek. Our next experiments studied the timing differences. Previously, we mentioned several times that quantitative genetic (polygenic or metric) traits (see Chapter 2. How Genes Vary in Fish Populations) were probably the most important traits in local adaptation. The allozyme marker documented timing differences, but it is not involved in determining timing. We needed to conduct breeding experiments to accomplish that. We conducted two experiments, one with families of early-run fish and one with families of late-run fish. Fish were tagged with tiny magnetic wire tags that were engraved with a unique code for each family. Magnetic tags can be detected externally in both juvenile and adult fish. The tags that we used were half the size of the tags that are normally used to mark other salmonids, because pink salmon fry are too small for the larger tags. Unsurprisingly, progeny of early-run fish returned early in the season and progeny of late-run fish returned late in the season. The critical observation, however, related to the nature of family returns within each of the experiments. Individuals returned at times closer to those of their relatives than to the times of unrelated fish, which means that there was a significant quantitative genetic (additive) contribution to return time. The implication is that there is variation in timing that can respond to natural selection. This suggests that return-timing is inherited within each spawning aggregation and is probably important to local adaptation, but that there is some year-to-year variation in the optimal timing for return.

The genetic marking experiment demonstrated that run timing has genetic determinants, and that the gene flow between them is restricted to a small number of individuals each generation. An obvious question is: “How does that work?” We mentioned above that Auke Creek pink salmon return in two pulses, one in mid to late August and the other in September. We speculated that the bimodal return was a result of habitat limitation. Regardless, during that time of year seasonal changes lower stream temperatures rapidly. During the two to three weeks that separate spawning peaks, the temperatures often decline from above 15°C to less than 12°C (Figure 4).

Following fertilization, the temperatures continue to decline until late November when they stabilize at between 0 and 3°C. The difference between early and late runs of 3°C does not seem large, but salmon are poikilotherms (cold blooded). The rate of many processes in cold blooded organisms (such
What Does Genetics Have to Do with It?

As metabolism (and other physiological processes) is regulated by the temperature of their environments. In fact, the rates of many physiological processes double with each 10°C increase. Of particular interest to us is the effect of temperature on rates of embryological development from fertilized eggs to fry (fish large enough to leave the stream). It turns out that development of embryos to particular events, such as formation of pigmented eyes (eyed eggs) or hatching time, can be predicted approximately by the temperature history of the embryo, that is, a combination of the length of time they develop and the temperature regime to which they are exposed. In fact, degree-days or accumulated temperature units (ATUs) are commonly used to roughly predict when hatching and other developmental events will occur. ATUs are tallied by taking the average temperature each day and adding all the temperatures together. For example, during three consecutive days on which average temperatures were 14°C, 13°C, and 15°C, an embryo will accumulate \((14 + 13 + 15 = 42)\) ATUs. A fertilized egg produced in mid to late August in 15°C water will accrue temperature units much faster than an egg fertilized in mid September. If the time separating them is 20 days, the early-run fish will have accumulated about 270 ATUs (20 days at a typical average of about 13.5°C during that time) before the late-run eggs are fertilized. After the late-run eggs are fertilized, both early- and late-run fish will accumulate temperature at the same rate.

Emigrant fry from early-run parents leave the stream a little earlier than fry from late-run parents, but the stream temperature is cold and the fry from late-run parents will not be able to catch up before they leave the stream. All told, early-run fry can accumulate about 180 more ATUs than late-run fry, which is about 15 to 20% more.

What we can conclude from that calculation is that the average environment experienced by developing early-run embryos is warmer than that of late-run embryos. And, because these are cold blooded organisms, the difference has to be compensated for in some way so that seaward migration timing is not compromised. Why is developmental compensation necessary? Food is most abundant in Auke Bay just after the spring bloom of phytoplankton, which are eaten by zooplankton, which are eaten by baby pink salmon (fry). The timing of the seaward migration of pink salmon fry must match the food abundance, or the fry will perish. Therefore, the seaward migration is linked both to the timing of their embryological development and to the timing of ecological events in Auke Bay. The events are cued to the fry by the seasonal day length (photoperiod) and freshwater temperatures. In order for fry of both Auke Creek pink salmon runs to enter salt water at an optimum time, the early- and late-run fry must enter salt water at about the same, even if the thermal regimes that they experience during development differ.

Figure 4. Cumulative returns of early- and late-run adult pink salmon to the Auke Creek weir in 1985, and cumulative emigration of fry descended from early- and late-run parents in spring of 1986. Estimates of the proportions of early- and late-run fish were estimated from daily samples collected at the weir. The total numbers that returned each day were used to expand the proportionate estimates. The curves are plotted from 5-day running averages.
These terms sound like names of heroes from Marvel comic books: “Holy humpies, Batman, it’s Genotype by Environment Interaction and Phenotypic Plasticity!” Actually, these terms describe how the average phenotypes of populations change in response to different environments. We mentioned previously that phenotypic plasticity describes the differences that can result from physiological adaptation to different environments. Plasticity would generally be shared by all organisms of a species. We can demonstrate it graphically as follows. Let’s see what average phenotypes result for each of several populations in (say) three different environments. We are using both phenotype and environment very broadly here, but you could think of size at six months of age for the phenotype and food abundance for the environment. If the phenotypic differences from the different environments were due entirely to plasticity, the changes would be predictable. If we connected the phenotypes resulting in a population at each different environment, the lines describing all the populations would be parallel (Figure 6, left). That means that the phenotypic differences between populations all respond in the same way in different environments. If the phenotypic responses to environmental changes were unpredictable, that means that we need to know about both the genotypes and environment to predict the changes. This is genotype by environment interaction, and the lines connecting phenotypes of one population across environments will intersect the lines of other populations (Figure 6, right). From a practical perspective, it also means that it will not be possible to predict what may result from bringing a wild population into a culture system. Phenotypic plasticity/genotype by environment also applies to families within a population.

**Figure 6.** A comparison of phenotypic plasticity and genotype by environment interaction. The phenotypes of five hypothetical populations are shown for three different environments (A, B, and C). The phenotypes (e.g., development rate) of a population are connected across environments to track their responses. Differences in phenotype that result from plasticity follow parallel lines. Differences that result from genotype by environment interaction show intersecting lines.
One compensation mechanism is that ATUs do not predict development rates perfectly. Development is a bit slower (relative to ATUs) at higher temperatures than would be predicted. However, since development rate is probably related to fitness because it determines how fast the yolk will be used and when the fry will be ready to migrate to sea, it seems likely that there is a genetic component involved in the differences in development timing that we observed between early and late runs. If that is so, it is also likely that development rate is a quantitative trait directed by the joint effect of many genes. Recall again that quantitative genetics involves the inheritance of complex traits that are influenced by many loci (see Chapter 2. How Genes Vary in Fish Populations). Recall also that breeding studies are used to estimate the proportion of phenotypic variability in a population that arises from genetic variation and the portion that is due to environmental variation. Also, the results of breeding studies are specific to the particular population and to the environment in which the experiment is conducted. In order to level the playing field, we conducted experiments in parallel. We used chillers and heaters to produce two incubation environments, one that mimicked long-term Auke Creek temperatures for early-run embryos and one that mimicked a temperature regime for late-run embryos. We split each family of early-run fish between the two temperature regimes and also split late-run families between the two temperature regimes. By following both early- and late-run families in both environments, we were able to compare development rates of the two runs and determine how they responded to the two different environments. We conducted experiments in both even- and odd-year broodlines. In those experiments, we separated the genetic and environmental components of development times (ATUs) to an early stage at which embryonic cells have completely surrounded the yolk mass, which is called epiboly, to the developmental stage at which the eye was first pigmented, and then to hatching.

We asked four questions. First, is there a genetic component in the expression of these traits? Second, do early and late runs differ in development rates in those three different developmental stages? Third, do differences that were observed between temperature regimes result from phenotypic plasticity or genotype by environment interaction (see sidebar 2)? And fourth, did the results differ between broodlines?

The first question was answered by statistically (analysis of variance) separating the components of phenotypic variance to estimate additive genetic variance as was described in Chapter 1 (Even Fish Obey Mendel’s Laws). Recall that the experiments require breeding experiments. In those experiments, we raised two families that shared a single sire (that is how geneticists refer to males in these crosses), and every family was kept in a separate compartment in an incubator. Periodically we made observations of a sample of fish from every compartment. We conducted the experiment with both early- and late-run families. We then compared members of families that shared a sire to each other and also to fish from families that had different sires. If fish from the same sire resemble each other more than they do fish from other sires, it is because they received their father’s genes. We observed that the sire’s genes generally contributed little to early development stages, but the contribution increased as development progressed. That observation is interesting and reasonable because the egg and yolk are preprogrammed by genes from the dam (how geneticists refer to females), and it takes time for the genes of the sire reach the same level of expression as the dam’s genes.

The next three questions can be answered entirely or in part by the graph in Figure 5, in which development rate (the rate at which ATUs increased) to landmark developmental stages is graphed for early and late runs in two different thermal environments. The plots include both even and odd broodlines. The graphs show that the development rates of early- and late-run fish differ in the same environments in nearly every instance. This means that the early and late runs differ genetically in composition for at least some of the genes involved with development timing. We also see significant environment by genotype interactions for all three stages in the odd brood year crosses, but only one of three in the even year crosses. Visually, the graphs of the even and odd brood years differ, which suggests that they differ with respect to the genetic information that determines development rate, and the statistical analysis (analysis of variance) confirms that the two broodlines differ significantly.
These experiments that examine development rates in the environments (or similar reconstructions of the environments) to which early and late pink salmon runs adapt demonstrate that there are genetic differences between the two runs that almost surely reflect local adaptation to the seasonal temperature differences. In addition, the comparisons of the two genetically isolated even- and odd-year broodlines show that they differ genetically from each other. The mechanism of local adaptation appears to be able to find more than one solution to the challenge presented it by the environment. One of the reasons is that the sequences of environments, which fluctuate between years, differ between the two broodlines, even if the overall average is the same. The other reason is that random drift plays a role in determining which genes are in the population.

Now we will extend our interest in population structure to a little broader geographic area and include a number of pink salmon streams that are in the Juneau area. When we conducted the original allozyme survey of pink salmon streams in the Juneau area prior to conducting our genetic marking experiment, we noticed that many of the streams were similar in length, water flows, gradients, and so forth. However, in the genetic marking experiment and the quantitative genetic experiments, we learned that there was population structure within Auke Creek that was based on timing and location. So, the next question we asked was how are the stream, spawning location within the stream, and spawning timing within the stream related to the genetic structure? By the mid 1980s, we had allozyme data for early- and late-run pink salmon and for intertidal- and upstream-spawning aggregations for several streams. With those data, we conducted an analysis that included 12 odd-year broodline and 6 even-year broodline pink salmon streams. Both those analyses, as well as analyses of data from Asian populations, showed us that return timing was the most important factor in genetic divergence, more important than the stream itself. That suggests that there is more exchange between streams during a run than between early and late runs within a stream. In addition, we detected no divergence between intertidal and upstream aggregations within a stream. However, see sidebar 1. Finally, there were differences between streams within a year. These differences are consistent with random drift (see Chapter 1. Even Fish Obey Mendel’s Laws), but they could be related to the way in which samples were taken—ordinarily samples from a stream were taken over several days, not continuously throughout the run.

From these experiments, we discovered that pink salmon can have a very fine-scaled population structure that depends on stream, timing, and possibly location within the stream. Of the Pacific salmon species, pink salmon are reputed, probably deservedly, to have the highest straying rate. Higher straying may be more common in pink salmon because, with some exceptions, pink salmon populations inhabit short coastal streams. Many of those streams are small and ephemeral, conditions that would foster higher straying rates. Our work demonstrated that at some times and in some places pink salmon home accurately and that such accuracy leads to opportunities for local adaptation. In comparison to pink salmon, we would expect that chum salmon might have stronger structure because of their larger sizes and frequent use of spawning sites that have upwelling water. Chinook, coho, and sockeye salmon (usually) have more rigid habitat requirements during their early lives because they often live and feed in fresh water before migrating to the sea. And when genetic structures of those species
are examined, they usually show greater divergence among populations.

**Summary**

Because of the enormous changes caused by the last Pleistocene glacial advance, many salmon populations are very young, in a geologic time frame. The oldest might be as old as 10,000 years (probably Upper Yukon River chum and chinook salmon), but many are only several hundreds of years old, and a few are only decades old. Many coastal streams are unstable and even ephemeral, which means that many salmon populations are also ephemeral.

Salmon return to natal streams to spawn, but they do not home perfectly, which means that there is opportunity for gene flow. Gene flow ties local populations together into metapopulations. Gene flow is important for maintaining genetic variability in small populations that may lose variability from random drift. It is also a source of recolonization in ephemeral streams.

At the same time as gene flow spreads variation among populations, local adaptation removes maladaptive alleles. In fact, a combination of homing and local adaptation can create very fine-scaled genetic structure among different spawning aggregations within a stream. Quantitative genetic traits (e.g., size, fecundity) are the primary targets of local adaptation and in spite of gene flow, very fine-scaled population structure may exist in small streams as a result of local adaptation.

Both gene flow and selection act simultaneously, and in tandem provide a regional stability in the presence of interannual fluctuations in climate. The genetic variation that is harbored by these and other metapopulations is the only insurance the species has against inevitable long-term changes.
The term aquaculture includes a variety of different methods to propagate fish and shellfish as well as some plant species. Hatcheries, where fish and shellfish are artificially spawned and where embryos and early life stages are kept, are a prominent feature of aquaculture. Both freshwater and marine species are included, and usually some aspects of the process involve human intervention and artificial environments. The primary reasons for aquaculture throughout the world are economics and the demand for high-quality protein and fat, but hatcheries are also used to address conservation problems. Because humans are involved and artificial habitats are often used, there will be pressures on the cultured populations that potentially alter the gene pool. Cultured populations can reduce resources that are available to the wild populations because the cultured populations consume them. If cultured populations and wild populations interbreed, wild gene pools can be altered. These interactions can be important concerns. In this chapter, we examine some of the intersections of genetics and hatcheries.
Two of the questions that began this set of articles had to do with the genetic effects of aquaculture on fish (see Chapter 1. Even Fish Obey Mendel’s Laws). Most aquaculture is accomplished by using hatcheries at some stage of production. The first question was, “Do hatcheries change the fish?” and its follow-up was “If they do, is that always bad?” The second question was, “Do hatchery fish harm wild stocks?” For much of the discussion in this chapter, we focus on salmon, but the general principles can be broadly applied to many other species.

**Purpose of hatcheries**

We are now ready to address those questions, but first we need to ask, “What is the purpose of a hatchery?” The shortest answer is, “To extend habitat beyond what is available naturally.” Think about it. Many hatcheries are used to increase the numbers of fish above their natural numbers. In those instances, it is obvious that either spawning or rearing habitat (or both) is insufficient to produce the desired number of fish. The increased numbers of fish can be used for put-and-take sport fishing, commercial harvests, conservation of depleted stocks, aquaculture, or other purposes. Whatever the reason, hatcheries supplement habitat. In hatcheries, the gametes are usually removed from the fish and combined to produce fertilized eggs, thereby removing normal mating choices that the fish themselves might make. Because the embryos are usually incubated under controlled (artificial) conditions, survivals are often much higher than would occur in nature where predation, ice scouring, floods, and other events reduce numbers. Some of those events may involve natural selection for processes that are important in natural production such as redd selection (redds are the “nests” in the gravel that fish—usually the female—excavate for their eggs) or tolerance of some environmental rigor. For species like coho, steelhead, sockeye, and chinook salmon that are usually fed and grown for substantial periods before they are released, the survival of juveniles is also generally higher in hatcheries than in nature.

**Intensive and extensive culture**

There are two categories of culture, **intensive** and **extensive**. Intensive culture raises organisms in captivity and (in theory) they are never released into nature, unless they escape. Pen-reared Atlantic salmon are an example of intensive culture. Most intensive culture relies on captive domesticated brood stock. In some instances, the hatchery maintains the breeders, and in others growth to market size is done at the hatchery from an early life stage (seed) provided by suppliers, who maintain the brood stock. In contrast, extensive culture takes advantage of the natural environment, such as the ocean or a lake, for at least a part of the cultured organism’s life cycle. Salmon ranching is an example of extensive culture; young salmon are released to sea where they grow and at maturity they return to their natal streams to perpetuate the cycle and to be harvested. Extensive culture can follow several different strategies or combinations of strategies: (1) hatcheries can be in remote locations, or they can release their young at sites far from wild populations of the same species and use returns or seed for brood stock; (2) hatcheries can be located near, but not on the same river system as, wild populations and take their brood stock from its returns; or (3) hatcheries can be located in the same system as a wild population and take brood stock from a mixture of hatchery and wild returns that escape harvest efforts. Fish that successfully make it past harvest activities are referred to as the “escapement” by fisheries scientists.

**Effects of intensive hatcheries**

Now let’s address one of the questions that brought us here, “Do hatcheries change the fish?” and “If they do, is that always bad?” Intensive hatcheries present several biological concerns. Some of the same concerns may also be issues with extensive hatcheries. The primary genetic concern is loss of genetic variation. Because maintaining brood stock can be expensive, brood-stock numbers maintained in hatcheries are often small, which means that only a small portion of the gene pool (a small sample size) may be used to perpetuate the stock. Random drift, which occurs as a result of small populations
(see Chapter 2: How Genes Vary in Populations), results in random fluctuations of allele frequencies from generation to generation. As a result of random drift, rarer alleles tend to be lost; and in persistently small populations, variation at many loci will eventually be lost. Also, small population sizes are often accompanied by an increase in inbreeding. Inbreeding can increase the incidence of the number of loci that are homozygous (have two copies of the same allele), including by random chance loci that are homozygous for unfavorable alleles (see sidebar 1).

**Altered gene pools**

One of the ways in which the gene pool can be diminished is by the spawning strategy used to produce hatchery fish. We cannot hope to replicate the mate selection that goes on in nature, but we can ensure that we maximize the numbers of parents (and their genes). One important practice is to use many males. Yes, it is possible for sperm taken from a single male to fertilize the eggs of 10 or more females. Why would one do that? It means that fewer fish would have to be handled. This approach might appear to increase efficiency, but the eventual cost in loss of genetic variation means that it is a false economy. The problem is that a single male would contribute half of the genes for all of the offspring and drastically reduce the genetic variation. A good rule is to use a one-to-one ratio of males to females. Another problem occurs when the sperm of multiple males is mixed together to fertilize eggs from multiple females. Several studies have shown that in such mixtures of sperm, the sperm of one male fertilizes a large proportion of the eggs, and sperm from other males may fertilize few or no eggs. In order to make sure that many males contribute, it is necessary to conduct single-pair (one male and one female) matings, or at least to use very small batches of fish.

Inadvertent selection can also alter gene pools of cultured species. In salmon culture there is a natural tendency to select large “pretty” fish for brood stock. However, being large or “pretty” does not necessarily translate to fitness in either the culture or natural environment. In addition, selection for similar phenotypes may restrict the gene pool. Another problem is that the numbers of returning fish cannot be predicted. As a result, there is a tendency to select early returns as breeders in order to assure that there will be full incubators. The result of selecting the early fish can result in advancing the return time. We saw previously that return timing is a trait that is important to local adaptation. One consequence of altering return timing is that other timing-related traits, such as emigration timing, may also be changed (see Chapter 5: Fish Population du Jour). One example of inadvertent selection is the chinook salmon stock in Minter Creek in Washington. Prior to the 1970s, the middle of the return time was in mid October. In a little more than three decades, the return timing has been advanced to late August, nearly two months earlier (Figure 2). Think of the seasonal environmental changes that take place between August and October in western Washington! We can all think of strategies to reduce this effect. The simplest would be to take breeders throughout the return period. The first problem is hatchery capacity and conservation limitations on the numbers of fish a hatchery is allowed to release. So, why not take the eggs throughout the return period, and decide after the run is over which ones to keep? Well, it is illegal in Washington to discard eggs (and there would probably be a public outcry, anyway) and conservation concerns prevent people from moving eggs to another hatchery. Here is an instance that laws designed to conserve a stock are actually the cause of the problem!

**Domestication selection**

Another genetic process that can alter hatchery stocks is domestication selection. Domestication selection is especially likely to occur in closed hatchery stocks—ones that fail to incorporate new genetic material on a regular basis. To understand how phenotypic and genetic changes can occur, we need only to recall our previous discussions of local adaptation. It should be obvious that the habitats provided by hatcheries often bear little resemblance to those encountered by the wild populations. The manner in which particular fish are chosen as breeders and the incubation environment experienced by the embryos match virtually none of the conditions of the wild environment. If the juveniles are fed hatchery food and reared in freshwater hatchery ponds before release, the differences are even greater. Because of
What Does Genetics Have to Do with It?

Sidebar 1

INBREEDING

Most people have heard of inbreeding and think of it in terms of offspring that result from matings between relatives. The closer the relationship between the mates, the more inbred the offspring. Many human cultures and religions have taboos about inbreeding and some cultures even have specific rules for who can and can't marry.

The concern has a genetic basis. Most species have numerous deleterious alleles on their chromosomes, that are rare in the population as a whole. In humans, many different genetic diseases result from mutated alleles, which produce defective enzymes. Examples of these diseases are Tay-Sachs (1 of 3,500 in the Ashkenazi Jewish population—about 1 in 29 individuals in that population carry the allele); phenylketonuria (1 of 10,000); and galactosemia (1 of 60,000 births among Caucasians). Tay-Sachs is untreatable; phenylketonuria can be treated by removing the amino acid phenylalanine from the diet; and galactosemia can be treated by removing galactose or sugars that include a galactose molecule, like lactose, from the diet (galactosemia is different from lactose intolerance, which is also the result of an inherited defective enzyme). Most humans carry several recessive (see Chapter 1. Even Fish Obey Mendel’s Laws) alleles that provide recipes for defective instructions for the protein or enzyme that they specify. Because each person carries two alleles that specify each enzyme (remember that we are diploid) and only one is defective, they are not affected by the disease. Also, because most defective alleles are rare, it is unlikely that two unrelated individuals will both carry alleles for the same defective enzymes. Regardless of how rare a deleterious allele is, however, when related individuals mate, the chance that an offspring will inherit two copies

Figure 1. Tracking a recessive allele in a mating between close relatives, in this instance, a brother and sister.
of a deleterious allele increases substantially. For instance in the mating of a brother and sister, the chances are 1 in 16 that a recessive allele will be homozygous in an offspring (Figure 1).

Let’s follow the deleterious gene that \(p_a\) carries in the figure. First we track alleles from parents to their children. Because both children receive an allele at a locus from each parent, they both have a chance to receive the same deleterious allele (\(a\) from \(p_a\) in the example). The chances that both receive the deleterious allele are 1 in 4 because the chances are \(\frac{1}{2}\) that the sister will inherit it and the chances are also \(\frac{1}{2}\) that the brother will inherit it. The chance that both brother and sister inherit the deleterious allele is \(\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}\). Now we look at the chances that both will pass the allele to a child. If they both carried the allele (which happens \(\frac{1}{4}\) of the time), the chance that the sister would pass the allele to the child is \(\frac{1}{2}\) (the other half of the time she would pass the normal allele). The brother would also pass the allele \(\frac{1}{2}\) of the time. So, the combined chance that both pass the allele is \(\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}\). But remember that only \(\frac{1}{4}\) of the time did both brother and sister receive the deleterious allele, so the overall chance that a child from a brother-sister cross receives two copies of a deleterious allele at a locus is 1 in 16. However, both partners in any mating probably carry several loci that are heterozygous (see Chapter 1) for deleterious recessive alleles. Unfortunate results will occur in a child if even one of its loci is homozygous for a deleterious allele.

Matings between more distantly related relations have lower chances for producing offspring that are homozygous for deleterious alleles at a locus, but inbreeding can build up in a population and the buildup is faster in a small population than in a large population. A number of experiments have been conducted to determine the effect of inbreeding. Harold Kinkaid looked at growth and survival in inbred cultured rainbow trout. Relative to controls (fish from the population that were not purposely inbred), fish derived from brother-sister matings declined 17.4% in numbers, 22.3% in weight as young fish at the size they were “planted” in a stream, and 36.6% in weight at that size, and 22.3% in weight at the time at which they would be caught by recreational fishermen. Two generations of brother-sister matings, which increased the extent of inbreeding another 50%, decreased the number of fish by 47.9%, weight at “planting” by 54.9%, and weight of catchable plants by 64.5%. In addition, the number of crippled fry increased in the inbred fish and their food conversion efficiency declined.

Similar experiments have been conducted in many species. There have been a few exceptions to this pattern, and the exceptions are usually for large species that have been through severe population-size bottlenecks (see Chapter 2: How Genes Vary in Fish Populations). Survival through such bottlenecks presumably resulted in part from purging most of the deleterious recessives from the populations. Large populations do not have that ability, and it is likely that more populations perish than successfully pass through the bottleneck.
What Does Genetics Have to Do with It?

the differences, there is an enormous likelihood that a hatchery population will adapt to the hatchery conditions and that the genes that produce more successful hatchery phenotypes will increase in abundance. This process is called domestication selection.

Is domestication selection bad? As we said in Chapter 1, it depends. If the goal is to have a population that thrives under culture conditions, particularly in an intensive aquacultural setting, the answer is no. In fact, fish that are more suited to culture conditions can be produced more economically, and domesticating the stock may be one of the goals of intensive culture. In other culture strategies in which the cultured product has little or no opportunity to interbreed with wild populations, the conclusion about the influence of domestication will depend on its effect on the suitability of the cultured product for both intensive and extensive brood stocks and market. In some instances, fish fed in culture can become very aggressive in their feeding habits. A voracious fish might be advantageous in a put-and-take sport fishery. In addition, some cultured fish that are released from salmon ranches return at different times or at different locations from wild populations and have little opportunity to interbreed or be harvested in mixtures with wild fish. This strategy can reduce both ecological and genetic interactions between hatchery and wild fish.

Disease

Now it is time to include the second question, “Do hatchery fish harm wild stocks?” Unfortunately, the history of intensively cultured organisms shows that all too frequently they escape. The first concern about escapees and intensively cultured fish is not a genetic issue—it is a disease issue. Under crowded culture conditions, diseases can break out. Contact with diseased cultured fish can infect wild populations. Disease issues are usually foremost in fish culturists’
minds because diseases are ordinarily quite obvious. In contrast, most genetics issues are not immediately apparent, both because they are not readily visible and because they may generally reduce survivals without symptoms that diseases usually show. In addition, the severity of the genetic problems may take two or more generations to develop, and an effect such as inbreeding may grow incrementally worse over time.

Interbreeding
The extent of our concerns for the genetic effects of escapees from intensively cultured organisms on wild populations depends on the relative numbers and frequency of escape incidents. If they are rare and the numbers of escaped fish are small relative to the number of wild fish, the effect of gene flow may be small and maladaptive genes would probably be purged rapidly by natural selection. Unfortunately, there are many incidences of farmed Atlantic salmon escaping in Europe, and the numbers of escapees are large. There is substantial evidence that those cultured fish have contributed large numbers of offspring to nearby wild stocks. Furthermore, the original brood stock for those cultured Atlantic salmon was not exclusively from local populations, so the hatchery gene pool was not a result of local adaptation (except to culture conditions). In Ireland, some wild Atlantic salmon populations have been entirely displaced by farmed escapees.

Outbreeding depression
Another possible genetic impact is hybridization between wild and hatchery fish. Hybrids between fish of different populations can cause outbreeding depression, which is a reduction in the survival and overall productivity of the wild population. That is, outbreeding depression results in a loss of fitness. We have documented outbreeding depression in hybrids between pink salmon (see sidebar 2), and it has been documented in many other species, including largemouth bass.

Competition with native species
The final concern for escapees from intensive culture is not a genetic issue, but an ecological issue. Many cultured species are farmed outside their normal ranges. Two examples are Asian oysters (see Chapter 7. The Lowdown on Frankenfish) and Atlantic salmon. The native range of Atlantic salmon is Europe and northeastern North America, but they are being farmed intensively in western North America and in South America. So many Atlantic salmon have escaped their pens in British Columbia that they are caught offshore occasionally by anglers and commercial fishermen. They and their young have been observed in streams of British Columbia. So far there is no evidence that farmed Atlantic salmon can complete their life cycle in that the young have returned to spawn, and most attempts over the past century to introduce Atlantic salmon outside of their natural range have failed. But it is not unlikely that as a result of persistent escapes, Atlantic salmon will eventually colonize native habitat of Pacific salmon. At that time, we will have to deal with an exotic species that competes with native Pacific salmon species. Because of this concern, the state of Washington recommended that only monosex fish (unable to reproduce) be farmed in Washington (http://www.wdfw.wa.gov/fish/atlantic/manage.htm).

Effects of extensive culture
Now let’s consider the potential impacts of extensive culture, such as salmon ranching, on wild populations. For species farmed outside of their native ranges, we have to consider the strong possibility that they will colonize nearby habitat, as the pink salmon and other Pacific salmon did in the Great Lakes (see Chapter 3. History of a Salmon Population). Farming diploid (fertile) Asian oysters along the eastern seaboard of North America could also create this problem in the future as it already had done on the west coast (see Chapter 7). Most salmon ranching, however, occurs within the normal range of the species under culture. Some of the genetic concerns about the hatchery brood stock itself are similar to those we discussed for intensive culture. In particular, many hatcheries were originally stocked with available brood stock from other hatcheries rather than local fish, so there is a risk of inbreeding and the translocated brood stock may not perform as well in a new environment.

What about the effect on wild populations? If the culture facility and release sites are geographically remote, and therefore genetically isolated from wild
Sidebar 2

OUTBREEDING DEPRESSION

In an agricultural context, we often hear about hybrid vigor. Hybrid vigor is the improvement in phenotypes that may occur in hybrids between genetically distinct strains, with improvement over the parental strains. This kind of effect is often observed, especially when inbred strains that have low amounts of variability are interbred. The enhanced characteristics result from the increased amount of genetic variability in the hybrid. The production of hybrid corn and tomatoes by interbreeding different inbred strains is a familiar example from modern agriculture. Hybrid vigor can result from interbreeding fish populations too, but not always.

What often gets left out of the descriptions of the wonders of hybrid vigor is that if the hybrids (F₁ —see Chapter 1, Even Fish Obey Mendel’s Laws) produce a second (F₂) generation, the phenotypes of the second generation are often inferior to those of the parental strains as well as to the F₁ hybrids. Indeed, hybrid tomatoes and corn are often not fertile; that’s how seed companies make money selling you their seeds. Decreased fitness (ability to contribute genes successfully to the next generation) that results from hybridization is called outbreeding depression.

Outbreeding depression between two genetically divergent populations can occur in two ways—ecological and genetic outbreeding depression. First, if they inhabit different environments they may have become locally adapted to those environments. Hybrids between the two populations will have a mixture (50:50) of the genes of the two populations. The genes that the hybrid carries may be inappropriate in both parental habitats. The hybrids might be most successful in an intermediate environment (although I think it is unlikely that we humans could construct an intermediate environment that takes into account all of the features that are important to a fish). This decline is ecological outbreeding depression; the genetic mechanism results largely from locus-by-locus or piece-by-piece contributions to the phenotype.

The second mechanism that generates outbreeding depression results from the simultaneous accumulation of alleles at many loci that have coordinated functions. The assembly of alleles combines their effects and, most important, they interact with each other to generate a favorable genotype. In this instance, the whole is greater than the sum of the parts because the interactions between alleles at different loci (geneticists refer to it as epistasis) are an essential component of the phenotype. The assembly of alleles is referred to as a coadapted genome. In the first generation of hybridization, each offspring receives one intact coadapted genome from each parent. However, in the second generation as a result of Mendel’s second law of segregation (see Chapter 1), the coadapted sets of alleles will be disrupted because the alleles segregate at random during gamete formation, and fitness may decline (Figure 3).

Because both first- and second-generation hybrids and controls must be examined to evaluate outbreeding depression in hybrids between populations, the experiments that are required to measure outbreeding depression can take a long time. We have conducted a series of experiments that involve pink salmon, which has for salmon a very short 2-year generation time. A single study that was replicated in both even- and odd-year broodlines takes seven years, six to conduct the experiment and obtain data and another year to analyze and report the data. We are in the midst of our third experiment. Why conduct the experiments? We want to examine the possible effects of translocated stocks on wild populations and the interactions between hatchery and wild populations. Why so many experiments? One reason is that the results are unpredictable. Another reason relates to the nature of the scientific method. What do we mean? Keep reading!

The scientific method uses observations to formulate a “falsifiable” hypothesis. That means that the hypothesis must be constructed so it can be disproved. For example, if we are comparing the allele frequency distributions of two populations, even if we think the populations differ, our null hypothesis is that the samples from the populations that we are comparing are actually subsamples from one large population that has a particular allele frequency distribution. If through our analysis we determine that it is improbable that the two samples came from the same underlying distribution, we conclude that they differ. If we do not see differences, we do not conclude that they are the same; we conclude only that we could not detect a difference. Possibly with more powerful tests, we might see differences. It is critical to realize that
just because we cannot delineate two populations by comparing them in some way, we are not justified in concluding that they are the same.

After we have constructed an appropriate hypothesis, we test it with appropriate experiments (or analyses). The challenge to scientists is that it is virtually impossible to “prove” a hypothesis. For example, consider a coin tossing experiment in which we want to determine if a coin is fair; that is, has equal chances of landing heads or tails. How do we go about testing it? First we have to formulate a hypothesis to test. Ordinarily we will hypothesize that it is fair (even if we have reason to think that it might not be) because that is a falsifiable hypothesis. Then we will test the hypothesis by tossing the coin some number of times and evaluating the results.

Before we examine the results, let’s compare the possible results of a fair coin with the possible results of two different unfair (biased) coins—one that on average lands heads 25% of the time and tails 75% of the time, and another that lands heads 45% of the time and tails 55% of the time. For the fair coin, it should be obvious that most experiments will result in heads between 3 and 7 times. Go ahead and try it yourself! Sometimes, you may see 8 or 2 heads and even 9 or 1 head. All 10 heads or all 10 tails will be rare, but will occur every once in a while (one in 1,024 for each result).

We can plot the chances of seeing each result (Figure 4). If we compare the possible results of the fair coin (called a distribution) to those of the strongly biased coin, we will see that their distributions overlap between about 3 and 5 heads. This means that we could not distinguish between the coins (with a reasonable chance of being

Figure 3. An analogy for genetic outbreeding depression. Genetic outbreeding depression results from disruption of coadapted alleles at multiple loci. Here two populations are represented by puzzles, the pieces of which represent alleles at different loci. The favorable interactions between alleles within a population are shown by the good fit of the pieces of the puzzle. Each individual in the population has two complete sets of the alleles because they are diploid. First generation (F1) hybrids receive one complete set of alleles from each puzzle and may not exhibit effects of outbreeding. In the second (F2) generation, however, each offspring receives two of each allele, regardless of the original parental source. The results may be a set of alleles that do not work well together.
correct) if those were our results for 10 tosses. If we compare the fair coin with the slightly biased (45:55) coin, we see that their outcomes overlap over most of the range. So what do we do? We increase the number of tosses! This experiment may take a bit more time if you do it yourself. The figure at the right (Figure 4) shows the sets of results (distributions) that we would expect if the coins were tossed 100 times. Notice that the results for the fair coin and strongly biased (25:75) coin differ substantially and overlap only at 36 to 39 heads, results that are unlikely for both coins. If we compare the fair coin with the slightly biased coin (45:55), however, we still see substantial overlaps. Only if we had less than 40 or more than 60 heads, would we be able to choose between the coins with reasonable confidence. However, that 5% edge can make a lot of money for Las Vegas. If we wanted to increase the power of our experiment to distinguish between a fair coin and the less biased coin, we might increase the sample size (number of tosses) to 1,000. For 1,000 tosses, there will still be reasonable doubt as to which coin you have if you toss 473, 474, or 475 heads; however, all three results are unlikely for both coins. But what about a fair coin with a coin biased to have a 49:51 ratio of heads to tails? And we can keep on narrowing the difference! We hope you can see that when we say fair coin, we have to specify just how fair to be able to test it.

Now back to using the scientific method and testing hypotheses. We saw that it can be difficult to demonstrate that a coin is fair (i.e., prove a hypothesis). However, we also saw that some outcomes of coin tossing experiments could indicate that a coin is unlikely to be fair. For example, a single result of 0 or 1 or of 9 or 10 heads in an experiment that makes ten tosses is an unlikely result for a fair coin, but for outcomes of between 3 and 7 heads we cannot tell. By increasing the number of tosses to 100, we saw that we could detect smaller levels of bias, but between about 41 and 48 heads we could not differentiate reliably between 45:55 (biased coin) and 50:50 (fair coin). For both experiments there were clearly results that would reliably allow us to rule out the hypothesis that the coin is fair, i.e., reject the hypothesis. But we can nearly always construct a hypothesis that is only slightly different from the null hypothesis. It may be difficult to resolve the two hypotheses, even with large sample sizes. The result is that it is very difficult to prove hypotheses, so most clever hypotheses that scientists test are ones that if disproved, leave only the alternative hypothesis; they are falsifiable. In the above example, we originally had reason to doubt if the coin was fair. The hypothesis that we constructed, however, was that it was fair. After the experiment, we could either reject the hypothesis and conclude that the coin is unfair or fail to reject the hypothesis. Beware! Failure to reject does not mean that the coin is fair, just that we did not see evidence of it. This is much like the criminal system in which a crook gets away with the crime because there is insufficient evidence that they committed the crime.

Gee, what a long-winded way to explain why we conducted three different outbreeding depression experiments with pink salmon. The bottom line is that the more genetically divergent two
populations are, the more likely it is that we can detect differences. We had been working with Auke Creek pink salmon, in part because pink salmon have a short two-year life cycle. Also, there are two genetically isolated populations in the same stream, but they return in different years, which makes experiments possible that involve genetically different pink salmon. Because we did not know at the outset what kind of results we might observe, we wanted to look first for outbreeding depression in the most divergent populations that we could produce. To make such hybrids, we learned how to cryopreserve semen (freeze sperm). We hybridized even- and odd-year broodline pink salmon by freezing sperm from one broodline, holding it for a year on liquid nitrogen, and using it to fertilize eggs from the other broodline. For this first experiment, we initially assumed that after we released the fry we would have very few returning adult hybrids because a similar experiment had been previously conducted in Canada by Withler and Morley1; they recovered few hybrids. Surprisingly, nearly equal numbers of hybrids and controls returned to spawn. This was so unexpected that we had not planned (did not have sufficient funding) for a second generation and were unable to follow through properly with those fish.

The successful return of hybrid fish prompted a pair of follow-up experiments (even-year males by odd-year females and odd-year females by even-year males), which we followed for two generations. In both experiments, the F1 hybrid and control fish had nearly identical survivals, which we measured as returns of adults to the Auke Creek weir. In contrast, the survival of the F2 hybrids was only about 70% of the F2 control returns. Our null (falsifiable) hypothesis was that there would be no difference in survivals of hybrid and control fish. We were unable to reject that hypothesis for the F2 generation, but we rejected that hypothesis resoundingly for both F2 returns and concluded that the F2 hybrids between broodlines of pink salmon had lower survivals than the controls. The probable reason for the difference is outbreeding depression. Also, because the difference occurred in the second generation, but was not apparent in the first generation, it is likely that the genetic outbreeding depression rather than ecological outbreeding depression is the cause. That is, crossing the two lineages disrupted their coadapted genomes.

Our first definitive experiment demonstrated that outbreeding depression can occur in salmon. However, natural hybridization between broodlines is very unlikely to occur in the modern marine environment (but see chapters 3. History of a Salmon Population, and 5. Fish Population du Jour, especially the description of transplanting pink salmon to the Great Lakes). As a result, we conducted a set of experiments that were a bit more realistic, especially for concerns about stock translocations, which have been commonplace throughout the world for most of the last century. In these experiments, we used females from Auke Creek and produced hybrids with males from Pillar Creek on Kodiak Island and controls with males from Auke Creek. We followed the hybrids and controls for two generations. In the first generation of the even-year broodline experiment (brood year 1996), the hybrid and control fish had similar survivals. In the second generation, the survival of hybrids was about 70% of the survival of controls. This pattern was similar to what we observed for crosses between broodlines, which was consistent with genetic outbreeding depression (disruption of coadapted genomes). Note that we are interpreting the results, but not stating that we “proved” them.

In the second (odd-broodline) experiment, we observed a reduction in survival of hybrids in both the first generation (hybrid survival was about 60% of the control survival) and the second generation (hybrid survival was about 75% of the control survival). The odd-broodline experiment suggests that both ecological and genetic outbreeding depression occurred.

So how critical to the “health” of the populations are the differences that we observed? If subsequent adult populations reflected the numbers of surviving progeny, the population would decline to 75% in one generation and to about 10% after eight generations if the effect persisted. There is no reason to think that recovery from genetic outbreeding depression would occur rapidly. It should be apparent that translocating stocks may be quite damaging to local populations of the same species, and the effect might persist for some time.

populations, it is unlikely that the released fish would have a substantial genetic effect on wild populations. This sort of hatchery has been termed “segregated” because the artificially spawned hatchery population is genetically segregated from wild populations (http://www.hatcheryreform.us has technical discussion documents that describe segregated and integrated hatcheries in detail).

**Ecological problems**

There can, however, be conservation issues. The way in which fisheries are conducted can be hugely important because the survival of hatchery fish from fertilized egg to release into the wild usually is much higher than for wild populations. Although increased survival is not directly a genetic issue, it can cause an ecological problem. The harvests of well-managed commercial fish stocks are restricted to the ability of spawning adults to produce adult offspring. As a result, one of the goals is to ensure that a sufficient number of fish escape the commercial fishery to return to their natal steams to spawn. The problem is that the higher survival of embryonic fish produced in the hatchery, as compared to survival of embryos that are produced naturally, means that the hatchery fish can be exploited in the fishery at a higher rate than the wild fish and still produce an adequate number of returns to seed the next generation in the hatchery.

So what’s the problem? If both hatchery and wild fish are harvested together, harvesting at a rate that is most efficient for hatchery fish will overfish the wild populations. The wild populations will be unable to sustain their numbers and will decline. On the other hand, if the returning fish are harvested at a rate appropriate for the wild populations, there will be excess numbers of returning hatchery fish. Unfortunately, the solution to the problem is not as simple as waiting until the fish return to the hatchery to harvest them because for many salmon species the flesh quality deteriorates rapidly as the salmon complete maturation, and they become unmarketable. So why not just let them die? First, there will be public outrage at the “wasted fish” (even though the reason is a legitimate conservation issue). Second, disposing of the carcasses poses a large economic problem. And no, in the remote locations of many hatcheries, there is as yet no economically feasible alternative use for the carcasses. Of course, one solution is to locate the hatchery where there are no wild stocks. Indeed, one of the strategies used in Southeast Alaska to increase chum salmon production has been to release young fish at remote sites where there are no (well, few) wild populations. The young fish can be imprinted to those sites by rearing them in seawater pens for a short period just prior to release. The fish return to those sites, where they are caught in fisheries designed to catch virtually all of them—so-called mop-up fisheries. A related strategy is to choose brood stock that return at a different time from the wild fish and to conduct separately managed fisheries. Timing differences would avoid the mixed stock fishery problem as well as potential genetic interactions, but the possibility of colonization would have to be monitored, and a brood stock that has appropriate timing might not be available. We will talk about yet another approach to the problem in a bit.

**Gene pool swamping**

Extensive culture that brings cultured and wild fish into contact can have problems other than passing disease from cultured to wild populations and possibly overfishing wild populations. If the culture facility is near wild populations, it is likely (inevitable) that some cultured fish will stray into the wild populations. If the cultured stock was developed from local populations, its gene pool may be preadapted for that area, and one would think that the effect of straying salmon interbreeding with wild salmon would be small. However, domestication selection may alter the hatchery gene pool by changing allele frequencies and those altered frequencies are what the wild population will receive. The impact of the strays depends on both their numbers and the extent of divergence between the cultured and wild populations. Regardless, if gene flow from the cultured population to the wild population is persistent and not countered by considerable gene flow in the other direction, the wild population will eventually be swamped by cultured genes. Even if the gene pool is not swamped, gene flow can move the genetic composition of the wild population away from the distribution of phenotypes that resulted from local adaptation. Some of the differences in phenotype can be subtle, but important, for example,
development rate (see Chapter 5. Fish Population du Jour). Coupled with different exploitation tolerances, this kind of hatchery system will often jeopardize wild populations.

**Balance upset**

Another extensive strategy is a hatchery that is located in the same stream as a wild population. The brood stock is taken from a mixture of hatchery and wild returns and the wild population includes spawners that started life in the hatchery (an “integrated” hatchery program in the terminology of http://www.hatcheryreform.us). Of course, potential disease issues remain, but genetic issues diminish if the number of cultured salmon returning to spawn in the wild is restricted and a substantial proportion of wild-spawned salmon are incorporated into the hatchery brood stock each generation. If the numbers of returns are properly balanced, natural selection and local adaptation will continue to act on the total population and the extent of domestication selection will be limited. Unfortunately, success at low levels of culturing often leads decision makers to increase culturing levels, which could throw the wild-cultured fish ratio out of balance. The “balance” that allows natural selection to predominate over domestication requires that the hatchery release relatively small numbers of salmon; and in the short term, people are usually interested in harvesting large releases.

**Habitat rehabilitation and artificial spawning channels**

Two other strategies increase the amount of “natural” habitat. One is rehabilitation of previously used habitat that was degraded by natural or human processes. This approach would be ideal, if the habitat were restorable. Unfortunately, habitat rehabilitation usually means that land use must be taken from one use, especially agriculture, in order to pass back to the fish. The other strategy is development of artificial spawning channels, which provide additional spawning habitat where rearing habitat is not limiting. Notable successes have been observed for sockeye salmon spawning channels in the Babine Lake system in British Columbia. The same genetic concerns exist for these kinds of fish enhancement as for other approaches if this “new” habitat is seeded by translocated stocks. However, if they are seeded naturally, the risks should be low. This approach also requires availability of appropriate land and water sources.

**Pitfalls of culturing for conservation**

A final topic that belongs in this section concerns the possible snafus that can result from well-intentioned efforts to resuscitate a population that is in decline; that is, cultural activities conducted for conservation purposes. There are many reasons that wild populations may decline. Among the most common reasons are overfishing, habitat loss, and climate fluctuations. The salmon populations of the Columbia River in the Pacific Northwest are poster candidates for conservation issues. Chinook, chum, sockeye, and coho salmon, as well as steelhead, were all once very abundant in the Columbia River system. However, as a result of societal decisions in the last century, many dams were constructed to generate energy and provide water for agriculture. Not only do the dams interfere with the passage of returning fish, but many of the spawning grounds were converted from a free-flowing river to a series of water impoundments, which are no longer suitable for spawning or rearing. In addition, much of the water was dedicated to agriculture, which further reduced the habitat. But wait, there’s more! Many Columbia River salmon spend their marine lives along the Pacific Ocean coast to the north well into Alaska waters. Those fish are subjected to a gauntlet of fisheries that begin in Alaska, continue into British Columbia and along the Washington coast, and into the Columbia River. Independent of the problems that dams created and overfishing, the overall salmon production in the Pacific Northwest has declined since the 1970s as a result of (apparently) cyclical climate changes. Recently, a transitory increase in production occurred, but production decreased again. For all of those reasons, many Columbia River salmon populations have been listed as threatened or endangered under federal law.

Now we can get to the issues that involve genetics and conservation efforts. One idea to address the conservation problems of depleted populations is to
use hatcheries to supplement natural production. The rationale is that it might be possible to resuscitate a population by culturing a portion of the population in a hatchery until it is released as a smolt (the life stage at which salmon leave fresh water to enter the ocean and change physiologically in preparation for a saltwater existence prior to emigrating). In concept, the idea is sound because hatcheries extend the habitat capable of producing smolts. However, the assumption that is made (often implicitly) is that the overall survival and return of adults will be increased by the supplementation and, after a few cycles, the system will be self-sustaining without the hatchery.

The left column in Figure 2 demonstrates that the population will remain at a small size if no steps are taken to enhance it. In that instance, its gene pool will remain intact, except for random drift (see Chapter 2. How Genes Vary in Fish Populations). Clearly, the natural habitat can only carry the depleted number. Supplementation efforts culture a portion of the wild returns. In a successful effort, the population size would increase and stay at the larger numbers, as is shown in the middle column of Figure 5.

There are two ways in which the assumption can go awry. The first is that if habitat is limiting, the population size will revert to the original low levels when hatchery production is stopped. The second problem is a harvest management problem. The public often has expectations that if there are increases in survival, much of that increase should be available to harvest. Both of these problems have genetic implications. The fraction of the endangered population brought into the hatchery can represent a large portion of the gene pool. Because the cultured portion has a much higher survival than the rest of the population, alleles from the cultured fish will increase in abundance relative to the genes in the rest of the population. If supplementation does not result in increased population numbers in the long term because natural habitat was not restored or because harvest keeps the population small, the rate of loss of genetic variation (a bottleneck effect) and rate of inbreeding will increase rapidly (see Chapter 2). This phenomenon is called the Ryman-Laikre effect (Figure 5). The consequence of heroic efforts to revive a population may be disastrous.

**Summary**

Hatcheries are usually deployed to address particular problems—the reasons can be conservation, economics, or both. The success of a hatchery at solving the problem must be continually evaluated, and when the problem has been reversed or it is clear that the hatchery is not successfully addressing the problem, there should be a means to terminate its operation.

Hatchery practices, both purposeful and inadvertent, can alter the genetic composition of a cultured stock. Some changes are beneficial, but others may not be. If the stock is cultured intensively and there are few escapees, it is unlikely that they will harm wild populations. Hatchery practices can also change the gene pool of an extensively cultured stock. Some of those changes, in particular inadvertent selection, may not be as benign, but may not be obvious for several generations.
Intensive hatcheries culture organisms for their entire life cycle and either hold their brood stock or purchase seed organisms. As long as they do not escape, they are benign. However, escapees can spread diseases and alter the gene pools of local populations of the same species. Intensive culture practiced outside the native range of the cultured species can result in exotic introductions.

Extensively cultured organisms can affect wild populations negatively. If large numbers of cultured organisms persistently stray into wild systems, the wild gene pool will soon be swamped. Wild populations are often strongly adapted to local conditions. Even within a small stream, there may be multiple spawning aggregations that have adapted to different environmental conditions. Finally, interaction between translocated fish (whether stocked or from an extensive culture facility) can be devastating to the wild population if outbreeding depression takes place.

Potential problems from wild-hatchery fish interactions can be reduced by using local populations as the source of brood stock and by releasing the fish at remote sites far from wild populations. Efforts should be made to monitor the influence of the hatchery releases on local wild populations. Ordinarily, such measures would require genetic and abundance surveys of local populations before the hatchery is initiated, and during the entire time that the hatchery is being operated.
One of the most controversial topics in genetics is genetic engineering. Organisms that have been modified genetically by introducing additional genes, often from other species, are referred to as transgenic organisms or genetically modified organisms (GMOs). The scientific objective of these gene transfers is to learn more about how genetics is expressed and how the genetic instructions are organized in organisms. The practical reason for producing transgenic organisms is to develop strains that have novel characteristics that often have economic value. Transgenic organisms have been produced in species that span the entire range of living organisms, including fishes. In this installment we will examine genetic engineering and how it can be done, the kinds of characters that can be altered, and some of the implications of developing transgenic organisms.
Genetically modified organisms (GMOs); genetic engineering! There, I’ve said it. Many people are extremely concerned about GMOs and regard them as somewhere between creations of Dr. Frankenstein and some evil mutant from outer space. They often envision creatures like the one in the movie *Alien*. The term GMO flips powerful emotional switches in many individuals. In this chapter, we consider the ways in which genetics can be used to alter characteristics (phenotypes: see Chapter 2. How Genes Vary in Fish Populations) of organisms. Clearly, there are mixed opinions about genetic engineering. Many opinions have some merit, but often the opinions are driven by emotion rather than fact, so we will also provide some facts as well as our opinions of the pros and cons. We will consider the pros first along with some of the underlying facts, but that does not mean that we are ignoring the cons because there are clearly serious issues that must be dealt with.

**The pros of genetic engineering**

Let’s step back and examine some facts. Genetic engineering in the broad sense involves manipulating the genetic composition of an organism to emphasize a trait that is in some way advantageous—to the population, society, industry, etc. Three common methods that are used for genetic engineering are genetic selection, ploidy manipulation, and gene transfer (also referred to as transgenics). For example, local adaptation, which we discussed previously in several different contexts, is “natural” genetic engineering. The genetic composition of a population changes in response to natural selection imposed by local environmental factors. Also, domestication selection (see Chapter 6. Genetics and Hatcheries) is usually inadvertent genetic engineering—the hatchery population changes in response to the conditions in which the fish are cultured.

**Genetic selection**

Genetic engineering has been practiced for tens of thousands of years. The term domestication that we used in the previous paragraph was not coined to describe hatchery fish; it describes the general process of taking a plant or animal from the wild, genetically (by breeding) removing many of the wild survival traits, and enhancing traits that we as humans think are useful or even aesthetically pleasing. Every agricultural crop, every breed of dog, cat, chicken—you name it—has been subjected to centuries and even millennia of genetic engineering. The principles of artificial selection, also known as selective breeding (see Chapter 2) work the same way in natural selection. When two individuals are mated, the phenotype of the offspring will tend to be more similar to the phenotypes of the parents than to those of other members of the population (on average) if there is a genetic basis for the trait. That means that if you breed larger, fluffier parents, the offspring will tend to be larger or fluffier, but only if there is a genetic basis for size and fluffiness. To see just how successful we have been, tune in a dog show on your television. Have you seen a hairless Chinese crested (Figure 1)? Clearly we have made substantial progress (sarcasm intended).

What we should not lose sight of is that genetic engineering through selective breeding has improved lives all over the world and has successfully staved off hunger in many, but unfortunately not yet all, regions. Agriculture has developed crops that are resistant to many diseases, increased their production and efficiency, and altered many species to satisfy our tastes (both gustatory and aesthetic). Parallel progress has been achieved with animal breeds. For example
in recent decades, the fat content of pork has been reduced (at our request) until it is now advertised truthfully as the other white meat (although many prefer the flavor of the older fattier varieties). All of these results come from selective breeding. Many of our domesticated animals and crops bear little resemblance to their wild ancestral sources.

**Ploidy manipulation**

“But,” you say, “that kind of engineering is ‘natural’ genetic engineering.” You are concerned by less “natural” engineering practices. Let’s take a look at another way in which we can genetically engineer an organism: ploidy manipulation. Most of the organisms we have considered so far are diploid; they carry two copies of information (alleles) for every gene (locus) (see Chapter 1. Even Fish Obey Mendel’s Laws). Diploid (2n) organisms are designed to produce haploid (n) gametes that restore the diploid number when fertilization occurred (i.e., the union of an egg and a sperm). Also, recall that the process of meiosis involves pairing of homologous chromosomes (chromosomes that carry the duplicates of the same alleles) during meiosis. It is part of the process that ensures that each gamete receives a complete haploid chromosome set.

One of the mechanisms that can isolate species is their chromosome complement. Most of the chromosomes carried by two closely related species have genes for the same traits and are very similar, if not identical. However, their genes may be arranged in different orders on their chromosomes, or the number of chromosomes may vary. For example, chum salmon carry 76 chromosomes (38 pairs), whereas pink salmon carry 52 (26 pairs). Natural hybrids between pink and chum salmon do occur. They are called chumpies, and carry 64 chromosomes (38 from the chum parent and 26 from the pink salmon parent). However, most chumpies are sterile because their chromosomes fail to find appropriate partners during meiosis (again refer to Chapter 1); and most of their gametes do not carry an intact haploid complement. As a result, mating hybrids to each other or to one of the parental species is usually unsuccessful. Just like a mule, the hybrid itself may be vigorous but it cannot produce offspring. Another example is a pink salmon egg fertilized by a chinook salmon sperm. The chinook-pink hybrids possess some characteristics that are advantageous for mass culture: they have early seawater tolerance like the pink salmon parent, carcass characteristics that are more similar its chinook salmon parent (most people would rather eat chinook salmon than pink salmon), and it is mostly sterile. In Alaska, the point is moot because penfish (intensive) farming is illegal, but in other areas that have numerous pink salmon and a few chinook salmon, the hybrid between chinook males and pink salmon females would be an attractive product. A few naturally occurring pink-chinook hybrids are found in the Great Lakes, but neither species is native there.

Induced polyploidy (see sidebar 1) can also be used to produce organisms that are usually sterile. Triploid females have been used in intensive (captive; see Chapter 6. Genetics and Hatcheries) trout culture because the females do not usually mature. This means that the energy ordinarily devoted to gamete production can be diverted to growth. Triploid fish have also been released to put-and-take fisheries. The idea is that they will not reproduce, but they will grow, so that the pond or lake will be populated by fish that grow to trophy size without contributing numerous small competitors that may prevent any fish from growing to appreciable size. Triploidy has been used extensively in the cultivated oyster industry. Triploid oysters provide a marketable product over most of the year, whereas diploid oysters undergo changes associated with reproduction that make them less palatable during summer months. Because of the triploid oyster, the industry in Washington state has grown tremendously since it was developed in the early 1980s.

What most people do not realize is that many of our agricultural crops are polyploids. The commercially cultivated banana is triploid (3n), has no fertile seeds, and is propagated from cuttings. Modern wheat is a hexaploid (6n) derived over thousands of years from hybrids that involved three different ancestral species. Most of our ornamental flowers and many of our cultivated crops are polyploids.

Polyploidy is a natural process. Salmonids (salmon and trout) and catostomids (suckers—the fish) descended from tetraploid ancestors. In fact, the entire vertebrate lineage was very probably made possible by two genome duplication events (see
THE WORLD OF POLYPLOIDS

There are many examples of organisms in which the entire chromosome set has been altered. Organisms that have multiple sets of chromosomes are called polyploids. On occasion, polyploids occur naturally. Of greatest relevance to us as vertebrates, are two chromosome doubling events that occurred during the emergence and evolution of the vertebrate lineage from invertebrates. There have been at least two events in which the entire complement of chromosomes doubled. The first event occurred at about the same time that craniate vertebrates (vertebrates with well defined heads) emerged. The additional genetic information may have provided the evolutionary fodder necessary to provide instructions for the new structures (head and gills) in lampreys and hagfish. A second doubling may have occurred at about the time that jawed vertebrates (cartilaginous and bony fish) emerged from the jawless vertebrates. It is likely that the extra information was modified to construct jaws and gill arches. After duplication, the genes and the chromosomes that carried them usually diverged until they became distinct—the tetraploid organisms became diploidized.

In addition to the duplications shared by all jawed vertebrates, some groups of species underwent a third duplication. Salmon and their relatives are descendants of a third duplication, and many of their genes retain a small amount of their tetraploid (4n) nature because the duplicated genes have not yet completely diverged. Members of the sucker family (Catostomidae) also arose through a ploidy event, but rather than a simple duplication of the chromosomes of an ancestral species, suckers arose from a hybrid ancestor in which the chromosome number doubled to include a complete set of chromosomes from each of the ancestral species.

These kinds of rearrangements were important in our evolutionary history, but they did not occur frequently. In contrast, polyploidy is very common in plants. Many cultivated plants, including numerous ornamental flower species, are polyploids. Another example is modern wheat, which arose from three different ancestral species, first through an ancient hybridization and doubling (to 4n) followed more recently by a second hybridization and doubling to a hexaploid (6n) set of chromosomes; hexaploid wheat undergoes normal meiosis.

When polyploidy results in an odd number of chromosome complements (3n, 5n, 7n, etc.) meiosis is disrupted. Most triploids (3n), pentaploids (5n), etc., are sterile because most species with odd ploidy are unable to produce gametes that have a single intact set of chromosomes, so there is no way that their gametes that will restore ploidy in the fertilized egg. Many cultivated plants have odd ploidy and only regenerate vegetatively. Often female triploids fail to produce gametes. For example, the common banana and seedless watermelons are triploid and boysenberries are heptaploid (7n); all are sterile and reproduce vegetatively.

Above we saw that polyploids can occur naturally and we looked at some examples. It is also possible to manipulate entire chromosome complements and “create” polyploids. The technology is simple. Heat, cold, or pressure shocks applied to recently fertilized eggs can block the second meiotic division. The result is that the egg provides two chromosome sets and the sperm provides one, and the result is a triploid.
The Lowdown on Frankenfish

Sidebar 1. The bottom line is that polyploidy is a natural phenomenon that has played an important role in the evolution of flora and fauna on earth.

**Transgenics**

The term “genetic engineering” often conjures up an image of mad scientists transplanting genes from one organism to another in exotic combinations that produce “monsters.” Although there is potential for abuse, there is also potential for societal benefits, such as enhanced world protein production and superior food quality. Because of the potential risks, transgenic research and product development in many countries are stringently regulated to prevent applications that might pose ecological, genetic, or health threats.

Although gene transfer is relatively easy to accomplish with plants, it is very difficult to do with animals. In principle, genetic selection accomplishes this slowly by aggregating the genes responsible for a desirable phenotype in a single population or lineage. Gene transfer, however, is more focused (sidebar 2). An authoritative source on the benefits and potential drawbacks of transgenic plants and world agriculture can be retrieved at [http://books.nap.edu/html/transgenic/](http://books.nap.edu/html/transgenic/). The publication was produced by the Royal Society of London, U.S. National Academy of Sciences, Brazilian Academy of Sciences, Chinese Academy of Sciences, Indian National Science Academy, Mexican Academy of Sciences, and Third World Academy of Sciences.

There are several experimental examples of transgenic fish. Genes coding for growth hormones have been successfully introduced into fish and expression of those genes has been documented. In many cases, faster growth has been observed. For example, a chinook salmon growth hormone gene was successfully introduced into chinook salmon, coho salmon, and rainbow trout, and an additional tilapia growth hormone was introduced into tilapia. A gene for an antifreeze protein, which is synthesized by some species of arctic and antarctic fishes, like the winter flounder, has been successfully introduced into Atlantic salmon (Salmo salar) with the hope that it might increase the survival of pen-reared fish in very cold waters.

Some transgenics may occur naturally. Dr. Peter Davies (Queen’s University, Ontario, Canada) reports that three distantly related species of very coldwater (−2ºC) fishes—herring, rainbow smelt, and sea ravens—naturally carry an antifreeze. Amazingly, the three antifreeze proteins are very similar and unlikely to have arisen independently in three parallel evolutionary lineages. Davies and his team speculate that the genes that specify the antifreeze proteins jumped from one species to another. Horizontal or lateral transfer (passing genes from one species to another) is common in bacteria, and occasionally viruses help genes move from one species to another. However, movement of genes from one animal to another has not been previously reported. The mechanism proposed by Davies might occur because most fishes have external fertilization. That is, males squirt sperm over eggs that have already been released or laid by females. If two fish are spawning at the same time and location, a sperm might stray and attach to the wrong egg. Ordinarily, a hybrid between two such different species would die. However, if the egg had already been fertilized by the right sperm, it might be possible for a small amount of the foreign DNA to be incorporated into the zygote. On rare occasions, the small amount of foreign DNA might include a useful gene, like the antifreeze gene. For more discussion see [http://www.theelectroniceconomist.com/science/displaystory.cfm?story_id=11703152](http://www.theelectroniceconomist.com/science/displaystory.cfm?story_id=11703152).

And now the cons of genetic engineering

The promise of genetic manipulation is the ability to design organisms with characteristics not found in nature, and others with characters that have been enhanced. This promise of genetic engineering also raises serious concerns. Intentionally engineered organisms may have a competitive advantage over native species and displace them or even disrupt the ecosystem. Because of the unpredictability and the potential adverse impacts, extreme caution must be exercised in developing and evaluating engineered fish.

**Genetic selection**

Natural selection is a normal process that occurs during local adaptation and evolution. The process is essential for the long-term persistence of most species. Both directed and domestication selection,
Gene transfer, or transgenics, provides another powerful tool with which favorable phenotypes can be developed. Of course, all tools should be used with care, but the development of molecular techniques has made genetic engineering possible and now provides exciting possibilities for developing aquacultural products. An entire industry has emerged that is based on genetically modified bacteria that produce valuable biochemicals. For example, the human insulin gene has been cloned into bacteria, which now produce insulin identical to that produced by humans for the treatment of diabetes. The process involves introducing a functional gene from another organism (a human in the case of human insulin) and getting the recipient organism to express that gene.

The technology for developing transgenic plants is quite advanced and relatively easy to accomplish, compared to animals. One of the reasons is that many plants (even sexual species) can reproduce vegetatively. The primary applications for transgenic crops at this time are herbicide tolerance and insect resistance. Transgenic crops that are grown in abundance worldwide are soybeans, cotton, corn, potatoes, rapeseed (canola oil), squash, and papayas. The United States, China, Argentina, and Canada are among the countries that grow these crops.

How is gene transfer done?
To accomplish the transfer, a gene (DNA) is introduced into the nucleus of a recipient cell; a successfully transferred gene becomes inserted into and becomes part of one of the recipient chromosomes. The gene transfer process is simple conceptually, but very difficult to accomplish in animals. Introduction is ordinarily done by microinjection (a miniature syringe-like instrument that punctures a single cell) or similar method to pass the gene across the cell membrane into the nucleus. Eggs or embryos are ordinarily used because there are only one or a few nuclei, the nuclei are accessible, and the tissue has not irreversibly differentiated into cell lines other than the germ line.

What are the requirements for a “successful” transfer? First, a successful gene transfer is permanent and the recipient passes its transgene to its offspring. Second, the offspring express the gene. Unfortunately, there is little control over where the gene is actually inserted into a chromosome; but expression of the introduced gene depends to some degree on the location at which it inserts into the recipient’s genome. Also, insertion of the gene into the recipient’s chromosome is no guarantee that it will be expressed, because gene expression is usually switched on by sequences that signal the position of the gene and the circumstances under which it is to be expressed. These signal sequences, appropriately called promoters, are often part of the transgene package that is transferred into the recipient. Expression of many genes is also enhanced or hindered by complex interactions with a variety of DNA-binding proteins called enhancement factors.
however, result in phenotypes that may not be “in harmony with nature.” In captive culture, the selected traits may increase economic value as a result of improved product or easier culture, such as fryer hens that are ready to market in six weeks. The phenotypes produced by directed selection may or may not be maladaptive in the wild, although chances are they will be, or else wild populations would evolve in the same direction. The cultivated fryer hens would not have survived long in my grandparent’s farm, where chickens roamed at will, because the fryer hens have adapted to being fed high quality (to chickens anyway) food and inoculated for disease. They are ill suited for making it in the nature of the chicken yard. There are other concerns with directed selection. One is that over many generations, genetic variation is generally reduced both by the selection process and by the relatively smaller numbers of breeders that are practical to maintain (see Chapter 6. Genetics and Hatcheries). A second concern is that a trait chosen for selection is often correlated to other traits that are also inadvertently selected for. For example, early growth rate in coho salmon is often correlated with early returning precocious males—jacks. This means that many stocks selected for rapid growth during their early lives will mature in a short time as very small individuals. It is well known that these precocious males are excluded from hatchery brood stock and have little economic value. Consequently, directed or inadvertent selection for early growth rate may be counterproductive in culture operations. In defense of these little guys, jacks often play an important role in maintaining genetic variation in populations. Obviously, they contribute genes that influence growth and maturity that may be less abundant in the rest of the population; but of more importance, they genetically tie together different brood years. For example, in many areas most coho salmon mature at three years. If the two-year-old jacks were not allowed to contribute genes, the populations would become three genetically isolated lineages (analogous to the two brood lines of pink salmon). Because the two-year-old coho jacks carry genes between what would otherwise be three genetically isolated lineages, they increase the genetic resources to all three, and help to stabilize the species genetically.

Most directed selection is designed for organisms produced by intensive culture; that is, pen culture or artificial culture systems that isolate the cultured organisms from wild organisms. Escapees from the culture facility may interbreed with wild fish, which can result in gene flow. As we saw in Chapter 6, occasional cultured strays may be genetically tolerated in a wild system because their few maladaptive genes will be removed by natural selection. However, persistent gene flow from the cultured stock will change the genetic composition of the recipient wild population. Note that pathology risks are a separate issue. Fish ranching operations (see Chapter 6), however, will probably produce persistent straying that can modify local wild populations. If domestication selection occurs in the culture environment, the genetic changes in the wild population are probably for the worst.

**Ploidy manipulation**

Clearly, there are some polyploids and hybrids that possess characteristics that may be useful or beneficial in a variety of fisheries management or cultural settings. Where a sterile or nearly sterile organism is advantageous—such as for a put-and-take fishery—they are ideal, particularly if the system is isolated (a pond or barrier lake) and the possibility of a very small amount of fertility is tolerable. Unfortunately, some interspecies hybrids are not completely sterile, and even offspring of hybrids like chumpies have a low survival (about 1%).

Under many circumstances where ploidy-manipulated organisms are cultured or even stocked, they pose little or no genetic threat in isolation. However, other species such as Asian oysters are often introduced for commercial culture outside their native range. One of the problems is that the process that produces triploids may not be 100% efficient and residual diploids may be present. Also, some triploids can undergo a bizarre meiosis that results in a few viable diploid gametes. Under such circumstances, the species may have sufficient fertility to establish self-perpetuating diploid populations, thereby resulting in an exotic introduction. Exotic introductions such as the rabbit to Australia or lampreys into the Great Lakes often cause severe ecological repercussions, but the extent and nature of their effects cannot be
predicted in advance. The Chesapeake Bay oyster industry has suffered severely from disease problems. The nonnative oyster from Asia, however, appears to do well there and is resistant to some of the diseases. The dilemma has been how to proceed—keep working with the disease-susceptible native species, culture triploid Asian oysters with a small risk of permanent introduction, or culture diploid Asian oysters that surely will take hold in the area. The conclusion reached by the National Academies committee (http://dels.nas.edu/dels/rpt_briefs/oyster_brief_final.pdf) was to cultivate triploid Asian oysters because introduction of the nonnative diploid Asian oysters would surely result in an invasion that has unpredictable consequences, and use of the native species would continue to cause economic duress and maverick farmers would probably introduce diploid Asian oysters on their own, anyway.

**Transgenics**

Gene transfers have the advantage that only one or a very few genes are introduced and need to be monitored. The disadvantage is that it is difficult or impossible to predict in which chromosome and at what location on the chromosome the gene will be inserted. Unfortunately, the site of insertion can influence the extent to which the introduced gene and genes near the insertion site are expressed. In addition, the promoters (see sidebar 2) that turn on expression of the transgenes require subcellular-level signals that will promote their expression. Questions that can result include, “Is the transgene expressed at appropriate times?” and, “Is the transgene expressed in a tissue that will make it useful?”

For plants, the issue is the effect of the transgenic product designed to ward off insects; in humans, it is the possibility of “shedding” genes; and in general the worry is the potential change in the ecological role of the transgenic organism. The most troubling issue for transgenic animals is the ecological effect they might have in the wild. It is difficult to foresee the possible negative results that might occur, but that does not mean there will not be any; and the possibility that a transgenic organism will behave like an introduced exotic cannot be ruled out without extensive evaluations. One modeling study suggests that a trait like size, which could increase mating success, might give the transgene an advantage in a wild system. The problem is that even if the progeny of the transgenic line were less viable, the transgene (which was referred to as a “Trojan” gene) might flourish in the population because of sexual selection, but ultimately erode the fitness of the population if the new phenotype (size for example) is maladaptive.

Recently, some experiments have examined the interactions of coho salmon that are transgenic for a growth hormone gene with wild coho salmon. Under normal fish culture conditions, the transgenic cohos grow very fast. They are voracious feeders and will eat the smaller wild fish, and eventually they dominate in the system. In experimental streams from which escape was not possible, the only food that was available was produced by the stream (e.g., insects). Under these conditions, relatively little food was available and the transgenic fish starved, which left only wild fish. This does not mean that releasing the transgenic fish poses no threat, but it does suggest that the threat may be less than we might imagine, at least in some instances. Obviously many more tests will need to be conducted before these fish would be considered safe for intrinsic culture applications.

**The first FDA-approved transgenic animal: good or bad?**

In September 2010, a transgenic Atlantic salmon that grows twice as fast as normal Atlantic salmon was approved for production by the U.S. Food and Drug Administration. A panel of experts reviewed the food safety and environmental risks posed by these salmon and concluded, “AquaBounty salmon was safe, as safe as food from conventional Atlantic salmon.” This is the first genetically engineered (GE) animal that will directly enter the U.S. food supply. Consequently, the process that was involved in its approval will serve as a precedent for future genetically engineered animals. Let’s examine the “product,” how and where it will be produced, and the concerns that are still being voiced about the decision. AquaBounty Technologies of Waltham, Massachusetts, engineered an Atlantic salmon that has an additional growth hormone gene (from the chinook salmon) and a gene that regulates its expression from the ocean pout (see sidebar 2). The ocean pout lives in subfreezing water for part of
the year and survives by synthesizing an antifreeze protein, which is expressed in cold water. This means that the chinook growth hormone is expressed in the winter, the GE Atlantic salmon will grow year-round, and, with the chinook growth hormone, it will grow twice as fast as normal Atlantic salmon.

Wait—what about the ecological concerns? Well, the AquaBounty hatchery on Prince Edward Island, Canada, will provide eggs for grow-out at other sites. That is where they expect to make their money—as an egg provider just like Monsanto provides sterile corn seed for farmers. Even better, the eggs produce triploid females, which as we have seen are mostly (>95%) sterile and presumably don’t include any males. It gets better. The only site approved for growing fish (so far) is in the highlands of Panama. In addition to physical containment, which should make it nearly fish-tight, a successful escapee would have to pass streams with temperatures much higher than normal Atlantic salmon can tolerate.

Why are there concerns, other than individuals who are concerned about everything? Well, this is the first approved GE animal, and because the process will undoubtedly serve as a precedent for future applications, it should be very carefully scrutinized. There are reasonable concerns about this particular application, but the primary criticism is that some of the supporting data and documentation were a bit sketchy. For example, no two fish culture facilities produce fish in exactly the same way or of exactly the same quality. For this application, the facility on Prince Edward Island and its fish were examined, but products from the Panama facility were unavailable. A second example is that the instrumentation used to measure the amounts of growth hormone and a suite of other hormones, including insulin-like growth factor 1 (IGF-1) lacked the sensitivity to detect biologically important concentrations. Why is this important? Production of IGF is stimulated by growth hormone. It is similar to insulin and involved in early childhood growth, has anabolic effects in adults, and most important has been implicated in cancer incidence. Finally, there was dissatisfaction over the extent of impact escapees might have on their new environments.

It seems clear that for such an important landmark—the approval of the first GE animal—the

Fact can be stranger than fiction, or Monsters of the Deep (and not so deep)

Mother Nature is the ultimate genetic engineer. Some of her results would chill the most enthusiastic monster film aficionado and her products even serve as models for film producers. Examples include human flesh-eaters, species that have absolutely disgusting habits, hideous looking creatures, and incredibly venomous fish.

The man-eaters

Of course, sharks immediately come to mind when we think of man-eaters. The most attacks on humans have been great white, tiger, and bull sharks (http://www.flnmh.ufl.edu/fish/sharks/statistics/species2.htm), although most of the Internet sources stress that the majority of attacks resulted from provocation or carelessness. Many sharks are indiscriminant about what goes into their mouths. A variety of weird (and indigestible) items have been found in the stomachs of tiger sharks that includes ladies’ pajamas, a roll of chicken wire, rubber tires, shoes, rags, a bag of potatoes, bottles, tar paper, a sack of coal, and a spam can (http://www.tiger-shark.info/). Clearly, they are not too choosy, so a wiggling appendage or two would not be ignored. At the voracious “killer Pomeranian” end of the spectrum is the notorious piranha. As you have doubtless seen many times on the screen, these fish attack in packs and can strip the meat from the bones of a large mammal in very short order. One of my friends kept an aquarium for piranhas—at the head of his bed. One morning he awoke to a piranha
lying next to him staring him in the face. I guess it miscalculated its jump.

**The ugly**

There are some frighteningly ugly fish out there that range from sort of cute to absolutely hideous. My favorite ugly cute fish is the spiny lumpsucker (Figure 2). These little guys are nearly spherical and move around rocks like little helicopters. I could watch one (and have, in an aquarium) for hours.

On the scarier side are the viperfishes (Figure 3). These fish live in very deep water and have bioluminescent structures on their bodies that probably attract food organisms and mates. Its fangs would make Dracula jealous, and are so long that they curve back along the side of its head. The hinged jaw opens very wide to trap unsuspecting prey. Their stomachs are large and expandable, so they can take advantage of prey when it is abundant.

**The gross**

Some species have disgusting personal habits. One is the South American candiru (Figure 4) or, as the locals call it, the willy fish. It is a small catfish. Candiru ordinarily parasitize the gills of larger fishes. They lurk in the mud at the bottom of the river and “sniff” out currents of nitrogenous wastes that flow from fish gills. Then they dart out, follow the stream, attach to the gills of the prey by lodging in place with its spines, and feed on the blood flowing through the gills. Occasionally they make a mistake and follow a current of human urine or other material, which also carries nitrogenous waste products, into a genital or other orifice where they are stuck because their spines prevent them from going backward. The consequences are bad for both the fish, which die, and the person they attacked, who suffer shock and infection. The fish must be removed surgically. Candiru can be surprisingly large—the largest grow to more than six inches. I would presume that the smaller ones are more dangerous to most humans.

Another misleadingly named fish is the pearlfish. Some pearlfish back into sea cumbers, tail first, and live on the cucumber’s internal organs. There is no getting around it; this must be uncomfortable for the host. The good news, however, is that some sea cucumbers expel their innards when threatened and then regenerate them. So, at least the host can replace its devoured parts.

The most disgusting of all fish, however, may be the hagfishes (slime eels) (Figure 5). These jawless marine fish lurk in the mud and wait for sick, dead, or hooked fish to sink near them. Hags then enter the fish through whatever opening may be available, even the eye socket, and eat the fish from the inside. Occasionally, their prey are hooked by fishermen. Just imagine on the deck of the boat a fish from which several “alien-like” creatures emerge. And just to be sure they retain the title of grossest conceivable fish, they secrete a slime that can turn a bucket of water into jelly in no time. So how do
they unslime themselves? They tie themselves into a simple (overhand) knot and pass the knot from one end to the other to remove the slime in a single glob. After one had slimed me and a laboratory sink, I was forbidden by a senior laboratory technician ever to bring a hag into her lab again. And she was tolerant of (although not an admirer of) lampreys!

The venomous/poisonous
A venomous organism delivers venom by injection or piercing. Poisonous organisms are harmful when eaten or touched. Many fishes use venoms or poisons as a defense. The poison tetrodotoxin shows up in Ian Fleming’s James Bond thriller, *From Russia with Love*. That novel ends in a cliffhanger. Bond was kicked by a spike-tipped shoe laced with tetrodotoxin, and we had to await the next book in the series to learn his fate. The neurotoxin is produced by symbiotic bacteria in the pufferfish (*fugu*) (Figure 6), which is the second most toxic animal in the world after the golden poison frog. Tetrodotoxin blocks nerve function. Tetrodotoxin is also concentrated from bacteria in the blue ring octopus, whose bite can be terminal. Tetrodotoxin is concentrated on the skin, liver, ovary, and other structures of the puffer. So how do puffers and other species that carry tetrodotoxin avoid poisoning themselves? Of course the answer is genetics! They have an alteration of a single nucleotide in the gene that encodes the structures (ion channel in nerve endings) that are blocked by tetrodotoxin.

Fugu meat and other parts are a delicacy in Japan, even though there are a number of deaths each year from fugu poisoning. Fugu chefs are licensed after they have received extensive training (10 years). The chefs are also expected to kill themselves if one of their customers dies from fugu poisoning. Raw fugu (sashimi) is very delicate and tasty. Yes, I survived several servings, but I did not experience the buzz that some say accompanies eating fugu. An interesting overview of fugu can be seen at http://www.thingsasian.com/stories-photos/3048.

The stonefish (Figure 7) is the most venomous fish in the world. Its dorsal fin includes numerous spines that release venom from two sacs attached to each spine. The venom is a mixture of proteins, which include hemolytic, neurotoxic, and cardioactive toxins. Typically, surviving victims suffer localized nerve damage that occasionally leads to atrophy of nearby muscles. The venom causes severe pain with possible shock, paralysis, and tissue death depending on the depth of the penetration. The pain has been called the worst pain known to man and victims often plead to have the affected limb amputated. A stonefish sting can be fatal to humans if they are not given medical attention within a couple of hours.

It should be clear from these examples and your own worldly experience that Mother Nature is an awesome genetic engineer.

Summary
Geneticists and culturists have an obligation to thoroughly evaluate potential genetic and ecological effects of introducing genetically engineered fish into natural systems or culturing them where they may escape. One of the fears many people have is that regulatory agencies will not be sufficiently stringent in their oversight. That is an issue that cannot be disregarded out-of-hand. But that particular fear pervades many areas of government oversight and is a political problem rather than a biological problem. Obviously, clear communication is part of the solution to such problems.

In spite of those and other caveats and issues, genetic engineering has the potential to increase protein and food production, which is badly needed to meet the world’s growing demand in face of dwindling resources. That fact alone will keep the genetic technologies in the foreground. Given that reality and the enormous economic potentials of transgenic organisms, we will need to apply our knowledge cautiously and wisely, during the next few decades. If we can accomplish that, we should be able to benefit from natural production and also take advantage of benefits that genetic engineering may provide, without compromising our natural resources.

Finally, it should be recognized that genetic engineering occurs naturally. Not only does Mother Nature produce some exciting and bizarre creatures, but some of the methods she uses might even involve transgenics, moving genes from one organism to another.
In this chapter, we examine how genetic methods are applied to stock identification. One of the objectives of fisheries management is to maintain stable, abundant populations that can sustain harvests over long time periods. In order to accomplish that, it is necessary to ensure that sufficient numbers escape fisheries to produce the next generation and that their spawning habitats are efficiently seeded. Although other phases of their life histories are also important, productivity of a species begins with reproduction. Spawning grounds that are too lightly seeded will not produce at their optimum, but overcrowded spawning grounds can also produce diminished numbers of progeny. Many species, such as salmon, are often harvested from mixtures of populations far from the natal streams to which they will return to reproduce. For these harvests, it is difficult to ensure that appropriate numbers of fish will eventually seed spawning grounds because there is often little control over or knowledge of which populations have been harvested. If the origins and/or destinations of the fish harvested in such mixtures could be determined, management would be substantially improved. This process is referred to as stock identification. Stock identification is also useful to establish allocations among user groups, to identify country of origin in order to satisfy requirements of some treaties between nations, and to determine the origins of fish caught incidentally as bycatch in other fisheries.
A challenge frequently posed to fisheries geneticists is, “If I get you a sample from the (say, eBay) fishery, can you tell me where the fish came from?” Often the question involves a species for which we have little population genetics knowledge. Consider a similar question in a different situation. M&Ms (the candy) come in many colors. If you were to mix several bags purchased from different grocery stores, would you be able to tell the store of origin of each piece? Of course not! If different stores sold different colors, though, it would be easy, but only if you had been to every store to determine the color that “marked” each store. Genetic markers work in the same way, but not quite as simply, because there are usually no population-(store-)specific markers. However, the abundances of alleles often vary between populations, which makes it possible to estimate contributions. We’ll continue with the M&Ms analogy. Let’s say that 75% of the M&Ms from Safeway are green and 25% are brown. In contrast, only 25% of the M&Ms from Albertsons are green, but 75% are brown. If you patronized only those stores, you would know that they were possible sources for candy in your bowl, and if 50% of the candy in that bowl were green and 50% were brown, it could easily be deduced that each store was the source of half of the candy (or that you had preferentially grazed on one or the other color—but we are not talking about selection here). Note that (1) we cannot unequivocally determine the origin of any single M&M, (2) we had to have previous (baseline) information for each store, and (3) we limited the possible contributors.

Now we’ll make the problem more complicated. My friend Bob also likes M&Ms and usually harvests a substantial portion of my stock when he visits. The last time he came over, he brought a large bag of M&Ms to supplement my bowl. It didn’t take me long to realize that he had purchased them at a different store because there were now red ones in the bowl. In order to determine the relative contributions from the different stores to the bowl (one of the burning questions of the day), I had to send out sampling crews to the other grocery stores in town to see which, if any, had red M&Ms. I could have asked him where he bought them, but I would still need to know if there were yet other stores that had red M&Ms for future contributions by generous guests. Also, I assumed that Bob did not bring them from “the outside,” because Bob is incapable of planning that far ahead. From our survey (I had lots of volunteers as long as it was at my expense), it turned out that only Krogers had red M&Ms; their bags held 50% red, 25% green, and 25% brown. After Bob’s visit, the bowl had 25% red, 37.5% green, and 37.5% brown. When I analyzed the stock contributions, it was clear (again assuming that no selection occurred during our grazing, and grazing was at random) that 50% of the M&Ms that remained after Bob’s visit came from Krogers, 25% came from Albertsons, and 25% came from Safeway. I get to eat them all now because who would want them after I have handled each and every one of them? If my bowl held mega quantities, I could have taken a random sample of M&Ms from the bowl and (assuming that they were well mixed) estimate the stock contributions from the sample. Because of sampling error (like we talked about for random drift; see Chapter 2. How Genes Vary in Fish Populations) it should be obvious that the larger my random sample was, the more accurately I could make the estimates.

The process of stock identification of fish from genetic markers is very similar to M&M stock determination, except that sampling is often not as much fun, and relatively small genetic differences among populations may make the exercise much more challenging. Fish do have two advantages, however. First, many different traits (loci) are available (like if the M&Ms had different shapes, scents, textures); and second, every fish has two (diploid) copies of most nuclear markers (genes). Diploid markers provide a powerful lever for prying out estimates of contributions of multiple stocks (see sidebar 1).

By now it should be clear that baseline information must be available for the populations that occur in a mixture before we can estimate their contributions. The frequencies of markers (alleles at loci) that are surveyed must differ among the populations. The more they differ, the better the stock separation algorithms will work. One obvious issue is that for some species it is probably impossible to obtain baseline samples for every potentially contributing population. Fortunately, populations in many geographic regions are genetically similar


**Sidebar 1**

**HARDY-WEINBERG EQUILIBRIUM AND GAMETIC DISEQUILIBRIUM**

Remember that the genetic composition of a population can be simplified as a gene pool, that is, the relative abundance (frequencies) of each allele at a locus. For example, at a microsatellite locus, a population may have three alleles: 133, 137, and 141 alleles. Remember that microsatellite alleles are specified (labeled) by their sizes (see Chapter 4, Molecular Tools for Population Genetics). In the gene pool of a population, the 133 allele might account for 20% of the alleles at that locus; and the 137 might be 50% of the total and the 141 allele 30%. This is a simple way to quantify the populations, but the individuals that actually make up the population are diploid—they each have two alleles at that locus. That means we could have six different diploid genotypes: 133/133, 133/137, 137/137, 133/141, 137/141, and 141/141. We can predict the frequencies of these genotypes from the allele frequencies if mating is random with respect to these alleles, and studies of many populations of many species suggest that mating is indeed close enough to random that we can take advantage of that idea. That idea also assumes that there are no disparities in allele frequencies between males and females, either. The result is that we can predict the genotypic frequencies probabilistically: the probability of a 133 gamete is 0.2 (20%); the probability of a 137 gamete is 0.5; and the probability of a 141 gamete is 0.3. We got those probabilities from the frequencies of occurrence of the alleles in the population.

The genotypes differ visibly, so we can count them. And simply by counting them we can calculate allele frequencies. We do this as follows. The chances of seeing a 133/133 homozygote in this population is the chance that the male gamete carried the 133 allele and the female gamete also carried the 133 allele. So, 20% of the organisms in the population received a 133 gamete from their father. Of that 20%, only 20% also received a 133 allele from their mother: $0.2 \times 0.2 = 0.04$. Only 4% of the population would be expected to be homozygous for the 133 allele (represented by the little square in the upper left corner of Figure 1). A heterozygote (e.g., 133/137) can result from a 133 egg being fertilized by a 137 sperm or a 137 egg fertilized by a 133 sperm. So, the chances of observing a 133/137 heterozygote is $(0.2 \times 0.5) + (0.5 \times 0.2) = 2 \times 0.2 \times 0.5 = 0.2$, which is represented by the top middle and left middle rectangles in Figure 1. The inherent circularity, that allele frequencies can be calculated from genotypic frequencies and genotypic frequencies can be calculated (counted) from allele frequencies, indicates that an equilibrium will be present. This is referred to as Hardy-Weinberg equilibrium.

The genotypic frequencies depend on the allele frequencies, but not in a linear way. What happens when individuals from two populations mix, such as in mixed-stock fisheries, is that the genotypic frequencies are not what would be predicted from random mating. Invariably, there are more homozygous types than would be predicted (Figure 2). We can see from the bar that represents the mixture that $38\% + 38\% = 76\%$ of the individuals are homozygous.

![Figure 1. Graphic display of Hardy-Weinberg frequencies that result from random mating within a population.](image-url)
What Does Genetics Have to Do with It?

From the allele frequencies in the mixture and Hardy-Weinberg predictions, we see that only 25% + 25% = 50% would be expected. Every mixture will have more homozygous types than would be predicted from random mating (Hardy-Weinberg)!

In addition to the increase in homozygous types at all loci, there are also differences in the multi-locus genotype frequencies. Think about it this way—if one population has two loci that have abundant alleles (say A at one locus and B at another locus, one would expect to see many individuals that carried both A and B alleles. If another population had an abundance of a and b alleles, it would have many individuals that had both of them. In a mixture of the two populations, you would expect to see an abundance of those combinations at the expense of individuals that had A-b and a-B combinations (Figure 3). This kind of disparity is called gametic disequilibrium because the alleles are quantified in pairs of one from each locus, just like gametes carry exactly one allele from each locus.

Stock identification programs use the underlying prediction of Hardy-Weinberg equilibrium and linkage equilibrium in each of the contributing populations as levers to pry apart (quantitatively) the disequilibria that exist in a stock mixture. To do so, however, they rely on the baseline data for the contributing populations.

**Figure 2.** Graphic display of the results of mixing individuals from two populations. Population 1 has frequencies of 0.2 for allele A and 0.8 for allele a, whereas population 2 has frequencies of 0.8 and 0.2 for alleles A and a, respectively. The phenotypic mixture is what would be observed in a 50:50 mixture on the fishing grounds. Random mating expectations would occur in a population with allele frequencies that are an average of the two populations, A = a = 0.5.

**Figure 3.** Graphic display of the results of linkage disequilibrium for two loci (locus A and locus B) from mixing individuals from two populations. Genotypic frequencies appear in each square. For population 1 and population 2, square colors reflect general abundances: red represents the highest abundance and dark blue the lowest. The sizes of the squares are not proportional to genotype abundances. HWE = Hardy Weinberg equilibrium.
to each other. The similarities could be a result of historic colonization, local gene flow (straying), shared similar environments (convergent selection), or a combination of those factors. Regardless, there is often substantial similarity, which means that we may be able to obtain good stock separation results by sampling a portion of the populations in each geographic region. Of course, we would have to verify the similarities. And, in the absence of strong divergence within the region, we might only be able to resolve our mixture to these regional groupings, rather than to specific streams; but in many instances, that level of delineation is sufficient.

Stock separation algorithms estimate the contribution of populations to a mixture by looking at the most likely way baseline samples could produce the mixture. Figure 4 shows, conceptually, how stock separation works. Once we have baseline information, we need to determine how well the data enable us to discriminate among populations or regional aggregates of populations. We can do this with simulations that use the stock separation algorithms to estimate contributions from synthetic mixtures generated by “in silico” sampling from the baseline. (Isn’t “in silico” a great term? It just means that the computer will generate a random data set that it draws from the allele frequencies that represent the baseline.)

Let’s say I wanted to see how well we could estimate the origins of M&Ms in my bowl. We would have the computer generate mixtures that we specify exactly ($x%$ from population or region A, $y%$ from population or region B, etc.) and then see how closely the stock separation algorithm comes to predicting the percentages of populations or regions ($x, y$) in the mixture that we had specified. For example, we could specify a mixture of 200 M&Ms, 50 from Safeway and 150 from Albertsons. The colors of the 50 M&Ms sampled one at a time from Safeway would be based on the baseline frequencies (75% green and 25% brown). This means that each M&M sampled from Safeway by the computer has a 75% chance of being green and a 25% chance of being brown; the process is like tossing a biased coin (see Chapter 6. Genetics and Hatcheries). The candies sampled from Albertsons would be sampled using its baseline information (25% green and 75% brown). Next we would use the stock separation program to estimate the contributions of the stores to these simulated mixtures. We would repeat this process many times (e.g., 20,000). Clearly, there will be random fluctuations in the estimates because they are random draws. However, if the stock separation algorithm works well, the average of the 20,000 iterations should be close to the 25% Safeway: 75% Albertsons that we specified. Also, if the markers are working well and we have adequate sample sizes, most of the estimates should be very close to that ratio; that is, there should be little spread around the true (input) value, which means that the variances of all of the estimates should be small.

Several factors contribute to the performance of the stock separation algorithms. The first is the extent of genetic difference among populations. Larger differences produce clearer separations. The second factor is the sizes of reference samples taken to generate the baseline. Large samples provide high confidence in the allele frequency estimates, whereas small samples may not clearly separate populations, which is analogous to the number of tosses used to test the “fairness” of a coin (see Chapter 6). The third factor is the size of the sample from the mixture. Remember that we could not determine the origin of a single brown M&M. So again, our confidence in the estimates of contributions improves as the size of the sample from the mixture increases.
One of the ways in which simulations can evaluate the baseline is to sample in silico (ooh! ooh! I get to use the term again) all of the individuals in a mixture from a single population or regional aggregation and see to where they get “misassigned,” for example, brown Albertsons M&Ms mistakenly attributed to Safeway (Figure 5). The process is repeated for each population or regional aggregate in turn. These simulations are very useful for identifying biases that result from similarities between populations (or aggregates of populations).

Now we will examine a specific example. For one of our projects, the ultimate goal is to learn about the marine distribution of immature chum salmon, particularly in the Bering Sea. Although a few Pacific salmon species spend their entire lives in fresh water (e.g., kokanee), most salmon spend substantial portions of their lives at sea; both Asian and Alaska fish commonly graze in the Bering Sea and some chum salmon from populations farther to the south also migrate through the Bering Sea. Recently, my colleagues A. Fuller, R. Riley, and S. Hall developed a preliminary baseline (work supported by the Bering Sea Fisherman’s Association) for chum salmon of data from microsatellite loci (See Chapter 4. Molecular Tools for Population Genetics). The baseline will be further developed, but here we will present a preview of what we expect the baseline to accomplish. First we will examine similarities among the populations represented in the baseline, then we will show the results of simulations that test the baseline, and finally we will use the baseline to estimate the composition of a sample of immature chum salmon caught in the Bering Sea.

First the baseline. We collected data for 12 microsatellite loci from samples of more than 4,900 adult fish from 73 populations that were distributed from the Pacific Northwest to Japan (Figure 6). We divided the populations into 26 geographic groups for analysis.

There are many ways that the allele frequency data of the populations can be reduced to provide
How Is Genetics Used for Stock Identification?

visual representations of the similarities of populations. Most people are familiar with trees (such as genealogies). Here we show another method called principal components analysis (PCA). PCA reduces the multidimensional data (one dimension for nearly every allele observed at all loci) to a very few dimensions that account for most of the genetic variation among populations. An analogy of this process is the use of maps. It is inconvenient to carry a 3-dimensional topographical map or marine chart, so cartographers reduce the information to two dimensions. Sometimes we are interested in different two dimensions, and that is what we see in a depth sounder while we are fishing. PCA reduces the multidimensional data for genetic variation to a smaller number (often two or three), but combines important information from several dimensions to generate drawings that often depict complicated data in a simple way. In this preliminary analysis, we reduced data from the 251 most variable alleles (251 dimensions!) to two dimensions (referred to as PC1 and PC2) to display the similarities among populations within a group and divergence among groups of populations (Figure 7). Most of the groups are distinct. Exceptions are the populations from coastal Western Alaska (Seward Peninsula [26-30], lower Yukon River [31-35], and Kuskokwim River [42-45]) as well as populations of Kamchatka (20-24) and the northern Okhotsk Sea (18-19) (numbers reference locations in Figure 6). We might expect that our simulations would be influenced by these similarities.

Our next step is to test the baseline by simulation to see how well we can delineate stock mixtures. Our baseline does not include every population in even one region. If we were addressing questions about populations within a region, we would need to have much more intensive sampling of the populations within the region. However, at this point we are broadly addressing origins of fish caught in the Bering Sea. Since the baseline shows strong regional similarities among populations, we will be satisfied (for now) with the baseline coverage that we have.

We conducted a series of simulations that evaluated how the fish from each geographic grouping would be assigned. To do this, we sampled in silico 200 fish from a single baseline region. If the differences among populations are sufficiently strong, we would expect that most or all of them would be assigned to the population of origin (as in Figure 5). However, genetic similarities between regions could result in misassignments. We repeated the
What Does Genetics Have to Do with It?

Simulation for every regional group (Figure 8). Our results showed that, although the assignments were not perfect, the majority of individuals sampled from a regional group were correctly assigned back to that group. The largest misallocations were among coastal Western Alaska populations and among southeastern Alaska and British Columbia populations, just as we suspected might happen from our PCA analysis. These simulations provide us with an indication of what biases might exist in analyses of mixed stocks.

Now we are ready to try out our baseline on a mixed fishery sample! In August 2005, a sample of 197 chum salmon was collected along the Aleutian Islands near Akutan. A second sample of 196 fish was sampled west of St. Matthew Island, which is about 700 km to the northwest of the Aleutian Islands site. We used two different programs to estimate the compositions of the samples. Both programs have strengths and weaknesses; and both results should be considered in interpreting the origins of fish in the mixture. The first analysis used the program SPAM, which we used in our simulations. The strength of SPAM is that it considers both the baseline and the mixtures as samples from the contributing populations and the mixture (which they are). What this means is that the program understands that the sample is not perfect. Samples do represent their sources; but since they are drawn at random (we hope) from the source, the estimates of allele frequencies may deviate a bit. (See Chapter 2. How Genes Vary in Fish Populations, which explains sampling error that results in random drift, and Chapter 6. Genetics and Hatcheries, which examines the variation in results of tossing coins.) For example, if we drew a handful of M&Ms from Safeway, we might actually have 4 brown and 10 green M&Ms, which is not exactly 25% brown (in fact, it is impossible to get 2½ brown ones!). SPAM uses random sampling procedures to account for these deviations and incorporates the uncertainty into the estimates of contribution of regions to the mixture (in the confidence estimates).
How Is Genetics Used for Stock Identification?

SPAM conducts the analysis from the perspective that every fish has an equal chance of having originated in any region. That is probably not a realistic assumption, however, because it is very unlikely that fish from all regions are evenly distributed across the Bering Sea. The second program (cBAYES) uses Bayesian analysis to estimate the relative contributions. Bayesian analyses incorporate other information into the analysis. For example, if you were invited to accompany a friend to a restaurant, you might be concerned about how to dress—suit and tie, business casual, etc. However, if I were the one who invited you, you would be reasonably safe dressing in business casual and probably showing me up. There is almost no way I would wear a tie.

In our stock identification example, many of the populations have the same common alleles at their loci. The result is that based on their genotypes, some of the fish could have come from several to many of the populations. SPAM would tend to assign those fish to the most likely of the highly likely population groups or partition them among the groups based on the likelihood. On the other hand, some of the alleles are prevalent in some population groups and not in others. Those genotypes would be assigned much more accurately. The Bayesian analysis makes a preliminary assignment

Figure 8. Simulations conducted to evaluate the baseline. The bar graphs are the results of 26 simulations, one for each geographic group of populations. In each simulation (individual stacked bar) the synthetic “mixture” was sampled from a single population group, and the contributions to the mixture were estimated using the software SPAM. The population group from which the sample was drawn and its fill pattern for that group is beneath its bar.
and then includes that information in subsequent assignments until the estimates stabilize. The Bayesian analysis uses information in the mixture itself to refine the assignments. For example, if the first cycle of estimates of M&M origins from many stores (including Safeway, Albertsons, and Krogers) suggested that there were more M&Ms from Albertsons than from Safeway, subsequent cycles would include that information in determining where green and brown M&Ms came from. Because the Bayesian approach is computer intensive, it does not incorporate the procedures that SPAM uses to account for sampling errors in the baseline and mixture. Clearly, what would be best is a combined approach, but one is not yet available, and it would be extremely computer intensive. Both analyses show substantial contributions from Asian and southern North American populations, from Bristol Bay and the Alaska Peninsula, and from coastal Western Alaska (Table 1). Few upper Yukon or Southcentral Alaska fish were detected.

These results tell us that chum salmon from Asia and Western Alaska, as well as from Southeast Alaska,
British Columbia, and even Washington, mixed in the waters just north of the Aleutian chain in August (at least that August). There were few fish, however, from the upper Yukon River and Southcentral Alaska at that time of year. In contrast, about 700 km to the northwest, most of the chum salmon in the sample originated in Asia. These results demonstrate that salmon stocks are not evenly distributed throughout the Bering Sea. In the near future, we will analyze many more samples to learn about the movements of chum salmon from season to season and also learn if climate differences between years alter their distributions. This kind of information should help managers and fishermen conserve our salmon resources and increase the accuracy with which we can forecast their abundances.

Remember that this is a preliminary glimpse! How can this application be improved? As the baseline continues to be expanded, the resolution should become sharper, and the results provided by the two types of computer analyses (or other analyses) should become closer and closer. Improvement of the baseline would include increasing the sample size of each population included. In addition, increasing the number of populations would improve the performance and increase resolution.

With a well developed baseline that had large samples of each population and extensive coverage, we would expect that the sampling error would decrease. Under those circumstances, the Bayesian approach would be the analysis of choice.

As mentioned above, if the objective were to refine estimates within a geographic area, we would need to sample thoroughly the streams within that region. If some regions or streams within a region remain difficult to resolve, additional loci will need to be evaluated. We are currently evaluating single nucleotide polymorphism (SNP) loci (see Chapter 4. Molecular Tools for Population Genetics) for inclusion in the baseline and expect that those loci will improve our resolution. The Alaska Department of Fish and Game is also developing SNP loci for chum salmon, which we may also eventually be able to incorporate.

Summary

Stock identification is an important fisheries management tool. The ability to estimate stock compositions from harvest mixtures enables managers to fine-tune harvests to ensure that optimal numbers of fish seed spawning grounds and that allocation of catches to user groups is appropriate. In addition, stock identification can be used to evaluate the impacts of salmon intercepted incidentally in other fisheries.

Genetics-based stock identification can be used for species whose populations are reproductively isolated from one another. The genetic compositions of isolated populations generally diverge over time as a result of random drift and adaptation to local selection regimes. Molecular genetic markers that result from the divergence can often differentiate populations (or geographic regions). Application of information from multiple loci often increases the resolution among populations (or regions). The stock identification process uses genetic information from potentially contributing populations (or regions) to estimate the likelihood that individuals or samples of individuals from a mixture originated from the reference populations. Ordinarily, the origins of an individual fish cannot be unequivocally determined; stock separation estimates are generally proportionate estimates of the composition of samples.

In this chapter, we explained the rationale for stock identification and provided examples of baseline development and validation. We also applied our chum salmon microsatellite baseline to the estimation of compositions of two samples of immature juveniles from the Bering Sea. A sample collected just north of the Aleutian Islands included fish from most of the chum salmon’s geographic range. A second sample collected about 700 km to the northwest included primarily Asian chum salmon. These samples indicate that there can be substantial variation in the geographic, and probably temporal, distribution of chum salmon in the Bering Sea. We are now analyzing other samples of immature chum salmon and expect to be able to learn much about the variation of spatial distributions of chum salmon stocks within and between years. Eventually, we expect to be able to correlate the variation in distributions of chum salmon with year-to-year environmental changes.
There are more than 100 species of rockfish (genus *Sebastes*) worldwide, most of which are distributed along the Pacific coast of North America. These fish are important contributors to commercial and recreational fisheries and critical elements of the marine food web. Although rockfishes occur in different colors and sizes, they are remarkably similar. Because there are so many species and because they change dramatically in shape from the time they are released as larvae by their mothers until they mature, they are often difficult to tell apart. Their similarity has also made it very difficult to figure out their evolutionary relationships. In this chapter, we will examine the methods that are used to deduce evolutionary relationships among species and use rockfishes as a focus. We will see that the process is not foolproof. Finally, we will learn that genetic information can be an important component of modern classification.
Like automobiles, rockfishes come in a variety of colors, makes, and models. There are more than 100 species worldwide. The Pacific waters of California are home to the largest number of species (about 65), but many can be found along the coasts of Oregon, Washington, British Columbia, and Alaska. A few species live in both North American and Asian waters of the North Pacific Ocean, and just over 20 other species live only along the Asian coast. During one of the warm periods between glacier advances (see Chapter 3. History of a Salmon Population), rockfish ventured from the waters of the North Pacific Ocean into the North Atlantic Ocean. And, during one of the very cold periods, other intrepid travelers moved from California to the southeast Pacific off Chile and thence to the South Atlantic Ocean. The rockfishes that now live in the north and south Atlantic diverged from those earlier colonizers into several distinct species.

Cataloging fishes

There are many reasons to develop a catalog of fishes. A practical reason is so that we can identify them. Why is that important? Fishes provide one of the important sources of protein for human consumption on this planet. In recent years, many stocks of fish have been depleted either by overharvest or by degradation of their habitats. In order to conserve those stocks and manage them at appropriate harvest levels, we need to know what we are harvesting. Many fisheries target a single species or a small number of species, but in other instances, especially in tropical and subtropical coastal fisheries, the target is all of the fishes that are caught. We will examine the interaction between genetics and harvest in a later chapter, but it should be obvious that if we cannot identify what we are harvesting, we cannot hope to conserve or manage fisheries effectively. In addition, both target and nontarget species are part of a very complex food web. Consequently, changes in abundance of one species can influence the abundances of many other species. In order to trace those relationships, we must be able to identify the components of the web accurately. A recent Reuters news release (http://www.reuters.com/article/scienceNews/idUSL2473263720080625?feedType=RSS&feedName=scienceNews) reports that the majority of animal species have not yet been described, although most of the nameless are not vertebrates. Moreover, many species have multiple names.

Another reason for cataloging species is to improve our knowledge of life on earth and the changes in species and their distributions that have occurred in their recent and distant pasts. That knowledge can give us insight into what may occur in the future. This purpose extends to addressing inherent human curiosity.

In order to keep track of the dizzying variety of species and to learn how to distinguish among them, we need a catalog or classification system. The practice of scientific classification is referred to as taxonomy or sometimes as systematics. For taxonomists (busily developing classification systems), the holy grail is to describe the relationships among species and groups of species. Another way to say this is: the goal of phylogenetics is to deduce the evolutionary relationships among species and to determine their origins. Often these phylogenies are depicted as trees. The ideal classification scheme would consist of discrete sets of species (or nested groups of sets, which we call taxonomic units or taxa—plural of taxon) that are monophyletic. Monophyletic means that members of a taxon share a common ancestor, which differs from common ancestors of other taxa. The same concept applies to taxa at all levels of classification; but, of course, as you move up the classification hierarchy to groups that include more and more species, the common ancestor would date back farther and farther. Species are classified into a hierarchy of levels like kingdom, phylum, class, etc. (see sidebar 1). And grouping taxa is accomplished by looking at morphological and, more recently, genetic similarities. The underlying assumption is that close relatives will more resemble each other than they will distant relatives. For example, a dachshund looks and behaves more like a wolf than it does a cat or a cow. Sometimes, the divergence is substantial; it is unlikely that a self-respecting wolf would claim a hairless Chinese crested mix (Figure 1) as a relative—even though it is. The wolf might eat it, but that is ecology, not genetics.
## Sidebar 1

### HIERARCHY OF FISH TAXONOMY

What we call a species and how it relates to other species is termed systematics. The scientific classification or systematics of fishes is organized hierarchically by breaking species into smaller and smaller groups. One of the important reasons for developing a systematic scheme is to provide biologists with a common means of communication. Let’s track a couple of familiar species—pink salmon and yelloweye rockfish—through the process one layer of hierarchy at a time. We’ll do it by examining a well-pruned tree of living fishes in which we include comments to help us track our species. This scheme uses J.S. Nelson (Fishes of the World, 4th edn., Wiley) as an authority. If you investigate phylogenetic schemes for fishes on the Internet, you will see conflicting organizations!

<table>
<thead>
<tr>
<th>Category</th>
<th>Group</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>Eukaryota</td>
<td>Fish are “higher” organisms, not bacteria (like E. coli and Staphylococcus aureus) nor archaea (a distinct lineage of single cell organisms that includes extremophiles like hot spring, thermal vent species, and septic tank methane producers).</td>
</tr>
<tr>
<td>Kingdom</td>
<td>Animalia</td>
<td>Fish are animals, which are distinguished from Protista (single-cell organisms that share the more complex cellular structure characteristic of animals—an organized nucleus), Plantae (plants), and Fungi (mushrooms, yeasts, and molds). Animals generally are motile and depend on other organisms (e.g., ultimately plants and members of other kingdoms) for energy.</td>
</tr>
<tr>
<td>Phylum</td>
<td>Chordata</td>
<td>Fish belong to the subgroup of animals that have a cartilaginous stiffener called a notochord and gill pouches or slits at some stage of their life. Other chordates are tunicates (Tunicata) and lancelets (Cephalochordata). You can look these up in Wikipedia on the web if you want more information!</td>
</tr>
<tr>
<td>Subphylum</td>
<td>Vertebrata (Craniata)</td>
<td>We separate vertebrates from the other chordates because they have an obvious brain enclosed by a head structure (cranium) and a spine (vertebral column).</td>
</tr>
<tr>
<td>Superclass</td>
<td>Gnathostomata</td>
<td>Fish, amphibians, mammals, and so on have jaws. Lampreys and hagfish (Agnatha) do not.</td>
</tr>
<tr>
<td>Class</td>
<td>Actinopterygii</td>
<td>Bony fish are separated from cartilaginous fish—sharks, rays, and chimaeras (Chondrichthyes), and the lobe-finned coelacanth and lungfishes (Sarcopterygii).</td>
</tr>
<tr>
<td>Order</td>
<td>Salmoniformes</td>
<td>Here is where our targets diverge. Pink salmon belong to a group of fishes that have no spines in their fins, no connection between their swim bladder and ear, and usually an adipose fin (that fleshy little flap on the back near the tail).</td>
</tr>
<tr>
<td>Order</td>
<td>Scorpaeniformes</td>
<td>The yelloweye rockfish belongs to the group of fish with a bone that extends from just below the eye to bones of the gill cover. Most are marine and many have venomous spines.</td>
</tr>
<tr>
<td>Family</td>
<td>Salmonidae</td>
<td>This group currently includes all of the living Salmoniformes.</td>
</tr>
<tr>
<td>Family</td>
<td>Scorpaenidae</td>
<td>This family usually has spines on its head and gill cover and venom glands at the base of some of the spines.</td>
</tr>
<tr>
<td>Genus</td>
<td>Oncorhynchus</td>
<td>This group includes the Pacific trouts and salmon and is distinct from the whitefishes (e.g., Coregonus), Atlantic trouts and salmon (Salmo), and others.</td>
</tr>
<tr>
<td>Genus</td>
<td>Sebastes</td>
<td>Rockfishes are live bearers.</td>
</tr>
<tr>
<td>Species</td>
<td>gorbuscha</td>
<td>Finally, we get to the pink salmon, which is distinguished from other species in the genus by its morphology. It can also be distinguished from genetic markers.</td>
</tr>
<tr>
<td>Species</td>
<td>ruberrimus</td>
<td>The yelloweye rockfish is also distinguished by color and shape. It can also be identified genetically.</td>
</tr>
</tbody>
</table>
What Does Genetics Have to Do with It?

Historically, ichthyologists (individuals who study fishes) used morphological characteristics such as coloration; the number, sizes, kinds, and locations of fin; the shape of cranial bones and the presence and locations of cranial spines; and the relative dimensions of the body to describe and to delineate species of fishes. They also used many of those morphological characteristics to try to deduce the evolutionary relationships among species. It should be obvious that not every characteristic that is useful for identifying and delineating species holds clues to evolutionary relationships. For example, wings emerged independently in insects, birds, and mammals (i.e., bats). This is called convergent evolution. But wings may not provide us with information as to which species are the nearest relatives. The character “possesses wings” is not restrictive enough for birds but is too restrictive for bats. Similarly, many car manufacturers produce both convertibles and SUVs (and some are both!). This convergent “evolution” in the automotive industry was driven by consumer demand (economics) and resulted in similarities between vehicles that descended from different automotive lineages. Similar parallel results can occur in fish species that are subjected to similar environmental pressures. However, car designers, unlike fish, can steal design ideas from other manufacturers.

The “rocky” beginnings of rockfish taxonomy

Classification of rockfishes has challenged ichthyologists for more than 200 years. The confusion began even before 1829, when the French zoologist Georges Cuvier gave the group of fishes that we now call rockfishes the name Sebastes. The genus Sebastes, which means magnificent or august, is presently assigned to the taxonomic order Scorpaeniformes (see sidebar 1). Also included in Scorpaeniformes are the venomous lion fishes as well as sculpins, greenlings, and the blackcod (also called sablefish or sable).

The beginning of rockfish classification was, well, rocky. It was not as smooth as having Cuvier say, “Let there be the genus Sebastes.” The father of taxonomy, Carl von Linné a.k.a. Carolus Linnaeus, was the first to introduce a rockfish formally. In 1758, he described a fish from the Mediterranean Sea that he named Perca marina, which was not a rockfish. Three years later, he gave the same name to a completely different fish from Norwegian waters. The Norwegian fish probably was a rockfish and is now commonly referred to as a redfish. In 1772, Peter Ascanius, a student of Linnaeus, described Perca norvegica (the golden redfish), which is also distributed in northeastern Atlantic waters. It was a rockfish, and probably the same one Linnaeus called P. marina. So what’s the problem? When Cuvier described rockfishes and coined the genus Sebastes 70 years later, he pointed out that Linnaeus’ Perca marina must apply to two distinct species because rockfishes are not found in the Mediterranean Sea. Secondly, the genus Perca is the one to which the freshwater yellow perch and walleye belong. That’s right; Linnaeus had also described another species, Perca scriba, in 1758. Both the original P. marinus and P. scriba are now considered a single species in the family Serranidae, which is in a different taxonomic order (Perciformes).
from rockfish, but the same order to which yellow perch are assigned (see sidebar 1).

Throughout the history of systematics, there have been many such chaotic instances. To a large extent, the confusion at the start of rockfish classification is understandable. In the early days of ichthyology, few species had been described, and reference specimens, when they existed, were not readily available. Linnaeus’ initial descriptions give us a glimpse of the difficulties that can be involved in classifying species. His fumbling also suggests that systematics is often not a precise science and that it might not be unusual for phylogenies (trees that demonstrate relationships), which may be generally accepted at one point in time, to change after additional specimens or species accumulate, when different characters are examined, or if the data are reinterpreted. Often the different arrangements are proposed by different individuals, but sometimes an individual changes his or her mind. In general, when sufficient data are unavailable, personality and opinion often play major roles in developing dogma. And dogma changes. Even now, there is no universally accepted phylogeny for fishes.

**Why the confusion?**

What kinds of problems lead to multiple phylogenies? First, there are nearly 25,000 living species of fishes, more than the total of all tetrapods (amphibians, birds, mammals, etc.) combined. And there are undoubtedly many undescribed species in remote (notably tropical Asia, Africa, and South America) and inaccessible (e.g., marine trenches) areas. Second, the species that exist now are descendants that diverged from older lineages, which themselves emerged in the often distant past. Where did fishes and vertebrates come from? See sidebar 2.

Since the first fishes emerged from their invertebrate ancestors, new species have evolved and diversified in response to changing environments and acquisitions of beneficial mutations, new combinations of genetic material, or even newly available genetic material. Also, many species disappeared as new and more competitive species appeared, like the Ford Falcon was replaced by the sportier Maverick and smaller Pinto. Sometimes the changes were gradual and it is possible to track the changes over time (like in the shape of the Mustang); other times lineages were completely replaced (Studebakers are no longer made), and sometimes representatives of species that branched from the main tree a long time ago have persisted (e.g., lampreys and hagfish). A third problem is that some lineages are broadly successful and rapidly diverge to fill available slots in their environments—ecological niches. When that kind of expansion occurs, new species may differ very little genetically but the morphological differences may range from subtle to substantial. An enormous number of cichlid species in Lake Malawi (check out http://malawicichlids.com) diverged to fill the many diverse habitats available there. Such closely related groups of species are referred to as a **species flock**. An ideal phylogeny traces two sister taxa to a single ancestral taxon. The shape of the phylogeny for cichlids and similar species flocks is many tips (species) that radiate from a single point (like spokes on a wheel) because the explosion of new species may not permit finer resolution. Taxonomists would much prefer branches that have only two limbs. Rockfishes have been described as an ancient species flock. That is, a large number of rockfish species that emerged some time ago (very roughly 5 million years) have since diverged, but in some instances, not by very much.

Notice that the key word we used above to describe how phylogenetic relationships are determined by ichthyologists was “deduce.” That means that the constructed phylogenies are no better than the quality of information available. We also noted (sidebar 2) that at times the information may be sparse and require interpretation. As soon as we introduce a term like “interpretation,” we open up the process to the likelihood of disagreement and controversy. Consequently, it is not unusual for different “authorities” to catalog species somewhat differently in the enormous dictionaries that are required for so many fishes. In some instances, a particular group may be named differently or be inserted into a different section of the catalog. In other words, there may be personality involved and just looking at the process should be interesting! Because there are so many rockfish species, *Sebastes* provides some particularly entertaining stories. Let’s examine a couple.
Sidebar 2

A BRIEF HISTORY OF VERTEBRATES

In 1995, the scientific journal *Nature* published an exciting report of fossils that were found in southern China. The fossils, which had the defining characteristics of vertebrates, were discovered in layers of rock that dated back about 545 million years and were much older than any vertebrate fossils that had been previously discovered elsewhere. Of course, this report sparked controversy because vertebrate fossils that old were completely unexpected and because fossils are usually incomplete and difficult to interpret. The controversy revolved around which potentially ancestral invertebrate group these new fossils most resembled. However, in 1999 more than 300 additional fossils from a similar taxon of about the same age were reported; many of those fossils were intact and they provided detailed information. This newly discovered (but very ancient) animal was named *Haikouella lanceolata* after the town Haikou near the discovery site. The overall picture that emerged from studies of those fossils was of a well developed vertebrate that had all the characteristics that define vertebrates: a longitudinal stiffener called a notochord (which is reduced to vertebral discs in humans), a hollow nerve cord that runs along the back above the notochord, an obvious brain structure, segmented muscles (recall the segments of meat that you see in a can of tuna or salmon), and gill arches. In fact, these creatures possessed some even more advanced features like a heart, the rudiments of a tail fin, and paired eyes. Another fossil species (*Haikouichthys ercaicunensis*) discovered near Haikouella lanceolata, which means that it has a similar age, appears to be even more advanced and related to modern lampreys, agnathan (jawless) vertebrates. These fossil discoveries indicate that the very first vertebrates emerged millions or tens of millions of years earlier than had been thought previously. We can deduce this because the fossils are the same age as the sedimentary rocks in which they were imbedded.

Like the very first ancestral vertebrates, ancient fossils of *Haikouella lanceolata* were fish. They had neither jaws nor paired fins and did not look like most present-day fish. Of course, none of those pioneer vertebrates lives today, but there is a trail of fossils that provides connections between those extinct early fish and living species. In addition, there is an enormous diversity of living fishes that we can examine for clues to their past. An obvious question is: “Where (or better who) did the first vertebrates come from?” and a related question is: “How do we know?”

It makes sense to start with the second question because its answer provides tools that we can use to address the question about “who.” One tool is comparative anatomy, a field that looks for resemblances in the anatomical features of living species and, to the extent possible, their similarities to fossil species. Another more recent approach is to compare the instruction manuals that direct the construction and operation of structures, functions, and even behaviors of fishes. As we learned previously, those manuals are encoded in the DNA sequences of living things (see chapters 1, Even Fish Obey Mendel’s Laws, and 4. Molecular Tools for Population Genetics). Both tools assume that new species arise from modifications in preexisting models, so we should be able to track evolutionary changes over time. Tracking, however, is not necessarily an easy task because the trail is complicated, and there may be ambiguous or misleading signs along the trail (like the wings of bats and birds). Clearly, the success in tracking lineages through the ages and determining the relationships among species (both living and extinct) depends on the wise choice of characters.

One of the challenges to paleontologists (individuals who study fossils) is that the fossil record is incomplete. Fossils come from carcasses that are covered soon after death (e.g., by mud or sand on a lake or ocean bottom) so that they are protected from scavengers. Chemical leaching and biochemical degradation processes usually leave only the hard parts, although in some instances fine-grained silt can record the texture of surfaces and remarkable detail can be preserved, such as in the southern China fossils. Over time, a series of additional, newer layers accumulate over them. Geochemical processes and pressure from the overlying layers eventually produces the fossils, which are imbedded in sedimentary (from the sediments that produce them) rock layers. The oldest layers are deepest, unless geological processes like uplift (mountain formation), volcanoes, earthquakes, glacier advances, or erosion from wind or water runoff rearranges or exposes them. Unfortunately for the paleontologists who excavate and study fossils, the geologic processes that expose many of the fossil beds
that they excavate also destroy innumerable others. And other fossil beds lie too far below the earth’s surface to be accessed. Another limitation is that because most fossil beds do not include large numbers of layers that span long time periods, they generally only provide snippets of life on earth at that place and time. Sites for really ancient fossils are scarce because they have had more time to be destroyed by natural processes.

The second tool involves examining the instruction books of the species. Recently, genetic technology has progressed to the stage that the information encoded in DNA sequences can be used to improve our ability to detect relationships among species. Of course, we do not have sequences from the ancient (extinct) species. However, if you picture those first species at the root of the vertebrate tree and the living species as tips of limbs on branches of the tree, it should be apparent that they are all connected. Moreover, just like for morphological evolution, the molecular evolution generally works by altering the instructions (mutations and rearrangements in the DNA) of the organism from which it was derived. Usually, the changes accrue one at a time, a change in a single nucleotide within a gene at a time. The differences increase over the generations, and accumulation of mutations can lead to divergence of species, but much of the information remains intact. In fact, the genetic (DNA sequence) differences between recently diverged species may be so subtle that they cannot be readily detected. What this means is that the extent of DNA sequence similarity between species is an indicator of the degree to which they are related, that is, how close together they are on a limb. Large differences might indicate that two species occupy different limbs or even branches of the tree.

The genetic approach is powerful and often provides perspectives not available from morphological data, but it must be used with caution for several reasons. First, the DNA sequences in different genes (and even in different parts of the same gene) can change at different rates in different parts of the tree. This means that, as for morphological traits, many characters (genes) should be surveyed and some judgment may be required in choosing genes. Second, as the path that connects two species gets longer (passes closer to the trunk of the tree), the differences between species may become more difficult to interpret. Another way to say this is that random changes also occur that may make it difficult to detect a signal (divergence) through the noise. Third, the nature of gene expression and the structure of the genome (e.g., how many chromosomes are there and to what extent genes occur on the same chromosomes in different species) must be taken into account when the DNA sequence data are analyzed; it can actually be a complicated process. Finally, because only small samples of the DNA sequences are available for most species, care must be taken to avoid biased interpretations, such as those of the blind monks who examined an elephant and came up with very different conclusions, each of which was based on particular part of the beast that they examined. The most accurate picture will result from the analysis of both morphological and genetic characters.

The question we were asking was “Who did the first vertebrates come from?” Obviously, we cannot know this, but we can look at the groups of living invertebrates to see which one shares some vertebrate characteristics, resembles early fossil vertebrates, and is most closely related to living vertebrates genetically. A small invertebrate called a lancelet (also called amphioxus or Branchiostoma) has a notochord, gill slits, segmented musculature, and a nerve tube that runs along the top of the notochord. They do not, however, have an obvious brain, just a little blister at the front end of the nerve tube and no obvious head structure. Lancelets burrow in the sand and filter the water through their gills to feed. Best of all, Haikouichthys ercaicunensis and the lancelet are similar looking (Figure 2).

A recent comparison (Nature 2008) of the DNA sequences of the lancelet with those of several other species (human, chicken, puffer fish, lamprey, and tunicate) corroborated the conclusion that the lancelet is the invertebrate group that is most closely related to the modern chordate line (Figure 3). The instruction manuals (genomes) of the lancelet and living vertebrates show that the manual for the lancelet is smaller than those of jawed vertebrates and probably too small to carry all of the information needed to construct and operate the more complex vertebrate. The evidence indicates that two separate events preceded the emergence of vertebrates that increased the size of the instruction manual—the entire genetic complement (genome) was doubled two times! Subsequently, one set of genes maintained the normal operation of the species, while the other sets accumulated mutations and was used “to
What Does Genetics Have to Do with It?

One notable result was duplications of a battery of genes referred to as Hox genes, which are clustered close together and responsible for directing the layout and floor plan for the developing embryo. The newly duplicated sets of Hox genes made it possible for the emerging vertebrate to construct more complicated structures. The lancelet has only a single set of Hox genes, but the jawless lampreys and hagfishes have at least two sets. Developmental biologists have traced the functions of both sets and determined that they govern different facets of construction of developing embryos.

The emergence of jawed species (sharks, fishes, mammals, etc.) also required a more detailed instruction manual. It is possible that the second genome duplication followed the appearance of jawless fishes (agnaths) but preceded the emergence of gnathostomes (jawed vertebrates like us). The additional sets of Hox genes appear to be involved in directing jaw construction.

**Figure 2.** Depictions of (above) the amphioxus and (below) an early vertebrate, *Haikouichthys ercaicunensis*. The latter is based on interpretation of fossils. Source: V. Gewin. 2005. Functional genomics thickens the biological plot. PLoS Biology 3(6):e219 doi:10.1371/journal.pbio.0030219.

**Figure 3.** Phylogenetic tree of relationships that focuses on the origins of chordates. The tree is based on a comparison of amino acid sequences (deduced from DNA sequences) from 1,090 genes. The long branches for sea squirt (tunicate) and larvacean (a free swimming relative of the tunicate) indicate high levels of amino acid substitution. Reprinted by permission from Macmillan Publishers Ltd.: Nature 453:1064-1072. N.H. Putnam et al. The amphioxus genome and the evolution of the chordate karyotype. Copyright 2008.
The taxonomic journey of dusky rockfish

Classification systems change over time. One of the common changes is that names given to a particular taxon may be elevated or demoted in hierarchical rank or be changed completely. Several rockfish species were described in the late 1800s and very early 1900s; each was assigned to a different genus. Two of those species, *Epinephelus ciliatus* (described in 1810 [or 1813 depending on which authority you agree with] by Tilesius) and *Perca variabilis* (in 1814 by Pallas) were from Russian collections that were caught along the Aleutian Islands. It appears that Cuvier (1829) united them as a single species as part of his description of *Sebastes*; and for most of the time since then, they have been lumped as a single species, *Sebastes ciliatus*. Recently (2005), however, the two original species were resurrected as distinct species, but now as members of *Sebastes*. The taxonomic journey on which they embarked took more than 200 years and traveled from distinct genera, to a single species, and finally to separate rockfish species within a genus: *Sebastes ciliatus* (dusky rockfish) and *Sebastes variabilis* (dark rockfish). Will the journey continue? Only time will tell.

Don’t let your personality get in your way

As taxonomists have attempted to split arrays of species into narrowly defined categories, they often included prefixes to the names like infraclass, superorder, and subgenus, and even invented new terms like tribe and series. For example, in a more detailed version of fish classification, the serranids and yellow perch mentioned above are actually placed in a superorder different from rockfishes (see sidebar 1).

New ichthyologists, like new administrators, often try to leave their mark by reorganization. One of the signs is that the labels for rockfish phylogeny have periodically oscillated between the genus and subgenus levels (the rationale is that with so many species, we obviously need to add an additional level of organization). And, names of species have converged and diverged like *Sebastes ciliatus* and *S. variabilis*.

For more than a century, ichthyologists have wrestled with rockfish taxonomy. It is often difficult to know which characteristics define important evolutionary changes. One set may be useful for delineating species (e.g., one car is a station wagon and the other one is a sedan), but they may not exhibit evolutionary landmarks (like the Mercedes logo, which marks all cars of the Mercedes lineage) for the fish as a monophyletic group.

Sometimes personality and territoriality of ichthyologists intruded into the process. As the numbers of known rockfishes increased from a handful of species in the mid 1800s to more than 100 in the mid 1900s, the number of genera to which they were assigned fluctuated up and down from one to two to several and to as many as 22. Only one genus is now recognized for the rockfishes, by most ichthyologists. Often the changes were the results of quibbling.

In the mid 1800s, William Ayres, a physician in San Francisco, purchased rockfish specimens from fish markets and formally described them to the California Academy of Sciences, which at the time was a group (sort of a club) that met regularly and discussed their observations. Ayres was sufficiently astute to recognize that among his market specimens were three new species (bocaccio, yelloweye, and China rockfish), that were related to the Atlantic redfish. He described them and placed them in the same *Sebastes* genus. During the 1850s several more rockfish species from the Pacific coast were added to *Sebastes*.

So far, so good. But a biologist named Theodore Gill (of whom early West Coast biologists had nothing good to say) made a close examination of bocaccio and noticed that they were a little different from other rockfish species: they were more elongate and had fewer head spines and small scales. Based on these differences, Gill (1861) proposed that a new genus be erected for bocaccio, *Sebastodes*. Ayres accepted this idea but modified it to specify the presence or absence of strong spination on the head. Then he divided the 11 known California species into two genera, *Sebastodes* and *Sebastes* (like the Atlantic redfishes). Gill (1862), however, contradicted that division and insisted that only the bocaccio belonged to *Sebastodes* and all the others (including one new species) probably belonged to his own new genus, *Sebastichthyes*, which differed from the Atlantic redfishes.
Ayres (1863) maintained his position and Gill (1864) retaliated snidely:

“The value of the characters used to distinguish the genera Sebastes, Sebastichthyes, and Sebastodes is now so generally conceded by scientific men, that it is unnecessary to further argue in their favor. I shall only remark that the combinations and distinctions of forms by Dr. Ayres are unnatural and violate all natural affinities. . . .”

Harsh! Let’s hear it for scientific objectivity.

Earlier in the paper Gill also said:

“But for the benefit of Dr. Ayres, who may doubt the value of the character [scale size], the opinion of Dr. Günther, whose authority he will scarcely [contradict], is [given as proof].”

Gill’s conclusion was to split his new genus Sebastichthyes into three genera (Sebastichthyes, Sebastosomus, and Sebastosomus). But the expert whose authority Gill invoked to try to embarrass Ayres took a shot at Gill and suggested that Gill should spend less time generating new genera and more time describing fish, that is, if he ever actually looked at them.

In 1880, David Starr Jordan and Charles Gilbert contributed some additional California species to bring the total to 20 (of which 19 are now accepted as distinct species, but some have different names!). In doing so, they made another revision and reduced the number of rockfish genera back to two: Sebastodes (only bocaccio) and Sebastichthyes (the rest), but subdivided Sebastichthyes into two subgenera (Sebastosomus and Sebastichthyes—yes, the same name for both genus and subgenus). In 1882, they reunited Sebastodes and Sebastichthyes into Sebastodes. But Jordan restored Sebastichthyes again in 1885.

There was another flurry of descriptions of eastern Pacific species in the 1880s and early 1890s. In their analysis of 29 species, Eigenmann and Beeson (1893) focused on cranial shape and structure and the presence and location of cranial spines; they concluded that there were eight genera of eastern Pacific rockfishes. Shortly thereafter, Frank Cramer (1895) reevaluated previous work, but focused on Eigenmann and Beeson’s efforts. In referring to the characters that Gill used to specify genera, Cramer said that “. . . all the generic characters which he assigned have proved worthless.” He also said that subsequent workers with knowledge of fish that Gill could not examine and more recently described species “. . . found it impossible to draw the lines of generic separation indicated by [Gill].” Because many distinctions at the species level were based on cranial characters, Cramer conducted a thorough comparative examination of cranial structures of rockfishes. Cramer provided detailed drawings of many skulls and concluded that because the primary character (the extent of joining of two cranial bones) used by Eigenmann and Beeson is too variable and is insufficiently discreet, it is unreliable. He also pointed out that structural traits often vary with age as do the head shapes and the locations and sizes of spines of free-swimming and bottom-dwelling species. Bottom dwellers tend to have vertically compressed heads and their eyes are closer together near the top, which optimizes their field of vision; they often have thick bone and large spines for protection. Fish that live off the bottom are generally more compressed laterally and their eyes are on the sides, which provides them vision above and below. They often have smoother heads, which are lighter and more streamlined. Cramer essentially said that the characters that had been used are useful for delineating species, but many of them occur in response to (genetic) adaptation to the habitats they occupied and could be convergent characters. Cramer concluded that the Pacific rockfishes all belonged to a single genus, Sebastodes. This view was embraced by Jordan and Evermann (1896) who divided Sebastodes into 13 subgenera.

Thirty years later, Jordan (1930), clearly old and crotchety, reconsidered and split Sebastodes into an astonishing 16 genera, although most researchers stuck to Sebastes and Sebastodes. Those genera and Sebastes (all this time considered a north Atlantic taxon) were reunited as Sebastes by Matsubara in 1943. However, North American taxonomists retained Sebastodes for Pacific species until 1970. One of the reasons the merger was finally made was because of genetic studies of rockfishes and other species from related, but non-rockfish, genera. In 1968, Henry Tsyuki and colleagues noted that the
The degree of difference between *Sebastodes* and *Sebastes* species was not more than they observed among *Sebastodes* species, but much less than was observed between rockfishes and other genera in the family. That’s right, a geneticist contributed to reuniting the two estranged branches of rockfishes. Recently, genetic studies by Rocha-Olivares suggest that North Atlantic species probably descended from the same lineage from which the Pacific ocean perch originated; clearly Atlantic and Pacific rockfishes are monophyletic! And you asked: What does genetics have to do with it?

**So what has genetics done for us recently?**

How can genetics contribute to clarifying taxonomic relationships? One of the challenges of selecting any character or set of characters for generating phylogenies is that it is difficult to choose characters that will delineate taxa near the tips of the tree, which are either not potentially reflections of adaptation (and potentially convergent) or just noise. Near the trunk of the animal tree we can identify characters like vertebral columns or developed brains (see sidebar 2), but near the periphery we often do not have such diagnostic characters. In addition, the environmental experience of an individual can contribute to the expression of some traits. For example, the number of vertebrae within some species is higher in more northern (colder) locations. Fortunately, because the instruction manual for each species (its DNA complement) is heritable and passes through generations with relatively little modification, the DNA sequence itself provides additional characters for comparing taxa. Of course, not every gene evolves at the same rate and we have already seen that there is variation within species. But by using a large sample of genes (many nucleotides sampled throughout the genome), it is possible to obtain excellent phylogenetic information. And, in fact, sequence data can provide phylogenetic information that cannot be reliably obtained by morphology alone.

When two species appear so similar that they are lumped as a single species (such as the two dusky rockfish species *Sebastes variabilis* and *S. ciliatus*),
they are called cryptic species. Genetics can help us delineate species and detect cryptic species. It turns out that the dusky rockfishes can be distinguished by morphology and where they live, but genetic differences (genes or alleles), which undoubtedly underlie the differences that delineate the two species, have not yet been discovered. Although such genes and alleles obviously exist, finding them is akin to finding a needle in a haystack.

Let’s look at a couple of examples of cryptic species that were detected in the last several years by genetic differences. In both examples, after surveying the genetic compositions of collections of individuals that had previously been labeled as a single species, it became obvious that two similar, but distinct, species existed in the same area but did not interbreed to any large extent. One of the cryptic species pairs was rougheye rockfish, which hid under the name *Sebastes aleutianus*, and the other were two species of vermillion rockfish (*Sebastes miniatus*). The rougheye rockfish pair is particularly notable because rougheye rockfish have been harvested commercially for more than a century, but no suggestion was made that there might be two species until the 1980s, and unequivocal confirmation was not demonstrated until recently.

Genetic differences between closely related species often provide genetic labels or markers with which to identify many species. Why not just look at the fish themselves and identify them from morphological differences? Some species, even those that are not considered cryptic, are very similar. In addition, some species-specific characteristics do not develop until later in life. For example, rockfishes are live bearers—they give birth to tiny swimming larvae, which look very much the same (Figure 4). However, there are multiple differences in mitochondrial DNA sequences of species, which can be used to identify many of the species and can be detected in larvae, although there are some small groups of species that cannot yet be separated genetically, like the dusky and dark rockfishes. Studies that combined morphology and genetics to identify Alaska larval rockfish concluded that most larvae cannot be identified from their morphology. In addition to not yet having developed the characteristic morphological attributes that characterize adults, juveniles and larvae may take on coloration patterns that provide them camouflage in their local environments. Larvae raised in a tank in a laboratory may not resemble wild larvae of the same species.

Many juvenile rockfish also are very similar looking and difficult or impossible to delineate, even by experts. Juvenile fish are further developed than the larvae and may exhibit some of the characteristics that are found in adults or other characteristics that enable some degree of separation. For instance, studies that combined genetics and morphology to identify young-of-the-year juveniles from the Gulf of Alaska and from the California Bight were able to identify most species when used in combination. Some species may have emerged so recently, however, that neither genetic differences nor morphological differences differentiate their larvae and juveniles.

**Current rockfish taxonomy is all sweetness and light, right?**

Now let’s look at some of the things that we swept under the rug as we went. First, we implied that Linneaus, Ascanius, and others appeared to have missed the assignment of rockfishes by a considerable margin. In fact, ichthyologists still have not reconciled the phylogenetic placement of perciform and scorpaeniform fishes. A single characteristic, a bone that projects backward toward the gill from the eye, termed a suborbital stay, unites the scorpaeniforms. One opinion is that the orders Perciformes, Scorpaeniformes, and two others (one of which includes the flatfishes) may be a single lineage and that Scorpaeniformes represents a group similar to the lineage from which the others emerged. Other authorities split these orders. So, Linneaus and Ascanius may not have been as far off as you were lead to believe, and they still are in good company.

Secondly, modern classification schemes do not even agree on the family to which rockfishes belong. Historically, they have been assigned to the family Scorpaenidae. However, some current schemes extract rockfishes and some of their very close relatives and place them in the family Sebastidae and elevate other taxa that are subfamilies under Scorpaenidae to the family level for the other groups. The taxon elevator is still at work!
Although tremendous strides have been made in deducing phylogenetic relationships that are based on some selected DNA sequence comparisons, especially within *Sebastes*, an unequivocal phylogeny awaits the acquisition and analysis of sequences from many additional genes. One of the obvious results of recent studies is that some of the rockfish assignments to subgenera that were based on morphological criteria were probably accurate and identify monophyletic groups, whereas others do not work well at all. In addition, although some strictly Asian species and some strictly western North American species were assigned to the same subgenera based on morphology, the genetic similarities among species from each geographic region and the divergence between species of the two regions strongly indicate that they are not monophyletic. By combining morphological and genetic criteria, it should be possible to develop a phylogeny that is widely accepted. However, it may always be difficult to disentangle the phylogeny of species within a species flock. In addition, it is not possible to obtain DNA sequence data from fossils, and many gaps in the fossil record may never be filled.

**Summary**

Taxonomy, or systematics, is the science of cataloging living things. A phylogenetic tree is one way of visualizing the catalog, and ideally the tree shows evolutionary relationships among species and among the different lineages. Historically, morphological characteristics were used to describe those relationships, but more recently genetic similarities have been included or used exclusively because the DNA sequences of organisms carry all of the information that underlies the structure, function, and even behaviors of species. Moreover, over time the sequences vary only at sites that accumulate the mutations that alter the functions of the products they specify. The challenge to taxonomists is to choose characteristics that track the evolutionary divergence that separates species, and not the evolutionary “noise.” The problem that occurs in choosing morphological differences is that some traits converge as a result of exposure to similar environments or emerge independently (e.g., wings) in different lineages. The challenge in choosing genetic characteristics is that many genes evolve at different rates and respond to different selective pressures. In this chapter, we saw that opinion often prevails when data are insufficient to allow clear conclusions. In addition, in order to describe modern phylogenies completely, we need to use the fossil record to find ancestors of today’s species. Because the fossil record is far from complete, it has not yet been possible to position all species phylogenetically. Questions remain about relationships, particularly among species at the tips of some branches and among many of the limbs that connect to the trunk of the tree. Rockfishes have a large number of species. By following bits and pieces of the history of rockfish taxonomy, we saw that the criteria applied by scientists and their perceptions of relationships varied remarkably in the last couple of centuries. We also saw that a phylogeny accepted by all taxonomists does not yet exist and that there are still newly discovered or newly described species sprouting from the tree. Many areas of taxonomy, for fishes and other organisms, will provide challenges for decades to come.
On the Pacific Coast alone there are more than 70 species of rockfish and a lot of them look really similar to one another, do a lot of the same things, and live in the same places. And what do we have to thank for these similarities? Genetics! A shared genetic history is to blame for the many similarities, and you can blame genetic differences when species are too similar to distinguish readily. As we have seen, the instructions encoded in DNA sequences are passed down the evolutionary tree, with an occasional mutation. In the DNA sequences lies the potential for an organism’s structure (shape), function (physiology and biochemistry), and behaviors. In this chapter we examine some of the diversity that comes from the genetic blueprints that rockfishes carry.
Our friends the rockfishes
Being a rockfish biologist is one of the more humbling experiences one can have. It’s right up there with trying to explain to your eight-year-old why the dollar has declined against the euro. It’s almost a truism that, after looking at photographs or videos of these fishes, or even when having one in hand, we spend a great deal of time murmuring “I just don’t know what it is” or the ever-popular “What do you think it is?” And this is to be expected because, oh my, there are a lot of rockfish species. To put this in perspective, if rockfishes were mountain lions, there would be (1) 17 species that eat deer and 14 others that consume ants, (2) some lions that live in the canopy of redwoods and others that hang out in the desert, (3) mountain lions that can barely get out of their own way and hang upside-down in caves and ones that gracefully glide over the forest canopy, (4) solitary and territorial species and ones that travel around in herds, and (5) cats that simply adore Proust and ones that favor Jacqueline Suzanne. And there would be a little tavern down in Sausalito where you could find 14 species in one room and all of them sitting in the same booth.

This very messy situation started way back in the Miocene, probably more than ten million years ago. Okay, we think it all started back in the Miocene, because those are the oldest rockfish fossils we have found. Unfortunately, those fossils are not of primitive rockfishes, because the extinct forms had all the verve and brio of modern species. What this means is that we don’t really know when the first proto-rockfish peered out from under a rock. We do believe that the early peering did not occur in the northeast Pacific. We know this because the closest relatives of the rockfishes, the rockfish fellow travelers in the genus Hozukius, live in the western Pacific. The current thinking is that the earliest rockfish was a spiny, squat bottom dweller in the western Pacific and that the sleeker and more spine-challenged species ultimately are derived from this form.

What is peculiar about all of this is that the rockfishes have been particularly successful (if success can be defined as simply dominating various marine habitats) after they crossed the Pacific. In the northwestern Pacific, the ancestral home of the genus, there are only about 30 species, less than half the number found off our shores. Indeed, in some habitats, for instance the deeper-water rocky reefs of the northeastern Pacific, rockfishes seem to have taken over in much the same way as a jingle for trans fat–laced burgers lodges in your head and pushes aside that nice symphony by Ralph Vaughan Williams.

Why rockfishes have been able to outcompete other fishes for these habitats is a matter of speculation. It may be, for instance, that the generalized body form of early rockfishes allowed them to take advantage of opportunities as they arose. This and the ability to speciate at the (geologically speaking, of course) drop of a hat propelled the group into one ecological niche after another. Another possibility is that, for reasons thus far unknown, early rockfishes along the northeastern Pacific found themselves in habitats that were relatively depauperate of competing species. Or, on the third hand, it’s all random chance and yet another example of some ineffable law of this universe.

But what is it about rockfishes in particular that allowed for (and for that matter still allows for) such robust speciation? It’s not enough for there to be an opportunity to slide into an ecological niche—there has to be something about rockfishes that makes such sliding (and this usually means the evolution of a new species) so facile. What about them makes speciating a veritable breeze? While we aren’t sure of the answers, a rockfish’s internal fertilization might be an important factor. With external fertilization (just broadcasting eggs and sperm into the water and hoping for the best) you can be a bit sloppy about your pickup lines, but with internal fertilization (where one has to get really close and do just the right things) males and females have to be in complete agreement about who is who. This means that potential mates have to look, sound (rockfish do make sounds), and smell (at least some species may produce pheromones) just right. With all of these having to go just right, it might be fairly easy for a group of fish to become isolated and quickly evolve a whole new set of singles bar moves. Take the vermillion rockfish (Figure 1), for example: males and females are colored and patterned differently during courtship. Female vermillion rockfishes are lighter and blotchier than the more intensely colored males.
Even today, this grandiloquent tendency to form new species seems to continue. Not only are there a number of species that are relative newcomers (e.g., pink and greenspotted rockfishes) that perhaps evolved in the last several tens of thousands of years, but there are also incipient species that seem to be in the process of separating as we wait and watch. As an example, stripetail rockfishes seem to be on the cusp of a northern and southern split and we wish them well. On the Pacific Coast, rockfish speciation often seems to occur along either depth gradients with one species tending to be shallower and the other deeper (e.g., blue and widow rockfishes), or along a north-south axis where one species has a more northern geographic range (e.g., rosethorn and pinkrose rockfishes). What is interesting is that closely related species often overlap in depth and geographic ranges and, at least to our eyes, do not seem to do things much differently.

Variety is the spice of rockfishes

For whatever reason, today rockfish lovers (the humans, not the fishes) find themselves with a veritable cornucopia of delights, as rockfishes of virtually every size, shape, color, and behavior inhabit our waters. Now, this is not to say that the look of a species has been thrown together in a haphazard way, like things in Dorothy's tornado-tossed house in the Wizard of Oz. Body shape, for instance, clearly reflects a species’ behavior. Fishes that live in caves and crevices and rarely move about, such as cowcod, rougheye, and yelloweye, look like they have been turned out on the same metaphysical lathe. Despite being only distantly related to each other, all have large and craggy heads, large dorsal spines, and relatively short dumpy bodies. Species that spend more of their time in the water column and are more likely to have an active lifestyle, such as blue, northern, widow, and yellowtail rockfishes, have more streamlined bodies with reduced or absent head spines. At the extreme, the semi-pelagic shortbelly rockfish, which forms large and mobile schools, is built like a sardine or mackerel.

Similarly, the extensive color palette that graces rockfish skins does make sense from an ecological perspective. Generally, shallow-water species (like blues, dusksies, quillbacks, and yellowtails) run to winter colors (e.g., black, blue, brown, and tan) while species that live in deeper waters usually sport warmer hues including lots of reds and oranges. This is not just another example of Mother Nature’s fashion sense, but rather the result of good old-fashioned natural selection (we had to get back to genetics here, didn’t we?). Red, orange, and yellow light (the longer wavelengths) penetrate only a short distance in seawater (about 30 feet for red and 150 feet for yellow) while shorter wavelengths (blues and violets) may penetrate 400 feet or more. A red-pigmented rockfish living in 10 feet of water looks red, while one living in 200 feet of water, well below where the red wavelength penetrates, appears black. For fishes that want to evade predators or hide from prey, being black might be quite useful. Of course, while we can understand why a fish might want to appear black, why doesn’t a rockfish just produce black pigment? Answer: we just don’t know.

And then there are color patterns. Midwater species are usually counter-shaded with darker backs and lighter bellies, which tends to make them invisible from both the top and the bottom. Many benthic rockfishes are disruptively marked and have lots of mottling and blotches to provide camouflage against the heterogeneous bottom. A group of four
closely related species—treefish (*Sebastes serriceps*),
tiger (*S. nigrocinctus*), redbanded (*S. babcocki*), and
flag (*S. rubrivinctus*) rockfishes—have bold vertical
bars (Figure 2). When seen under natural light in
their usual rocky haunts, these fish look like zebras in
tall grass. Interestingly, as pelagic juveniles, all four
vertically barred species are often found in drifting
kelp and other debris. And to make things more
confusing, young splitnose rockfish, which as adults
have little marking at all, also have vertical bars when
they live in drifting kelp. This is not to say that we are
able to nicely pin the natural selection tail on every
color pattern donkey because sometimes rockfishes
are patterned just because, well, they are. Squarespot
rockfish, for instance, often have square spots on
their sides. The sides of halfbandeds are creased by
several dark brown bands, while starry rockfish are
profusely spotted with tiny yellow dots and the Asian
goma-sui (*Sebastes nivosus*) has gorgeous iridescent
blue spots against a dark background (Figure 3.).
Sharpchin, pygmy, and bank rockfishes have a “<”
mark that emanates from the back of each eye.
An interesting pair of Asian species (usu-mebaru
(*Sebastes thompsoni*) and *S. joyneri* (togotto-mebaru)
has similar distributions of large pigmentation patches
(Figure 4). However, one has nearly square patches
and the other more rounded patches.

Many species are capable of changing their colors
or patterns. Like many other fishes, stressed rockfishes
tend to become darker and more heavily pigmented.
This means that a fish yanked from the water by
an overzealous fisherman may look very different
from one at rest in the ocean. Some species also
routinely change appearance when underwater. This
is particularly true of such midwater taxa as bocaccio,
chilipepper, halfbanded, and squarespot rockfishes.
When in the water column, all of these species are
fairly monochromatic (with the exception of a few
specific markings on halfbandeds and squarespots).
However, within a few seconds of settling to the
bottom, all of these fishes acquire a series of quite
characteristic blotches and vermiculations. While it is
easy to assume that these disruptive markings make
the fish harder to see when it nestles among rocks,
we have seen this behavior in 1,200 feet of water,
where there is not enough light to make a difference.
Interestingly, this typical “benthic” patterning is very similar to the juvenile patterning of all of these species.

Over the years, these swell complications have made rockfish taxonomy a bracing and thrilling competitive sport. Our friend, Lo-Chai Chen, who described a number of very closely related species in the days before genetic testing, was once asked how he first distinguished species. “Gestalt” was his answer. And gestalt does neatly sum up how many species are first discerned (see Chapter 9. The Challenges of Taxonomy as Seen Through the Eyes of Rockfish Biologists). There are often subtle differences that someone, a fisherman for instance, who has seen thousands of fish, unconsciously keys in on. One of us was for a time a commercial fisherman in southern California. Other fishermen made him keenly aware that there were two forms of vermilion rockfishes: an inshore form, dubbed the “mainland reds” and a deeper-water one, the “island reds.” It was only recently that DNA testing demonstrated that there are indeed two closely related species, the inshore vermilion rockfish and the deeper-water sunset rockfish. Of course, and particularly in the days before DNA testing, a rockfish biologist’s personality was as important a factor in discerning a new species as any constellation of physical characteristics. As an example, Julius Phillips, one of the most knowledgeable rockfish researchers, was loath to describe any new species. Indeed, in his groundbreaking 1957 publication A Review of the Rockfishes of California, Phillips noted in passing that he had found three hybrids in his studies. All of these are now accepted as distinct species. On the other hand, Carl Hubbs, a man who once thought a fish covered in sauce that he was served for dinner in a La Paz restaurant was a new species, described a number rockfishes that have not withstood closer scrutiny.

**X(treme) rockfishes**

The focus of this chapter is variation. We’ve considered coloration, but what other interesting factoids can we come up with? What about size? The largest reported rockfish was a 120 cm (nearly 4 feet) shortraker rockfish (*Sebastes borealis*). Recently, a slightly smaller version (44 inches and 60 pounds) was captured by a pollock catcher-processor (Figure 5). This fish reminds one of the character “Fat Bastard” in the Austin Powers movie. In contrast, the largest recorded dwarf-red rockfish (*S. rufianus*) was about 6.75 inches, but only two of that species have ever been knowingly observed. The smallest commonly seen species is the Puget Sound rockfish (*Sebastes emphaeus*), which grows to a whopping 7.2 inches (Figure 6).

Some rockfishes are old enough to have voted for John Quincy Adams, assuming that a rockfish was in favor of the enlarging of the Dismal Swamp Canal in North Carolina. One shortraker rockfish was aged at 205 years when it was caught. Life spans of 50 to 100 years are not unusual. On the other end of the spectrum, the oldest calico rockfish (*Sebastes dalli*)
What Does Genetics Have to Do with It?

along with a couple other species, rarely live past 12 years.

What about depth? Anyone can hang out at the surface, but how deep can they go? A rougheye (or was it a shortraker) rockfish was pulled up from about 2,830 meters (about 9,300 feet)! When it was caught, biologists did not yet realize that rougheye and shortrakers are two distinct species.

Another amazing fact is that most females produce tens to hundreds of thousands of eggs per year, with a maximum of 2,600,000 estimated for a vermilion rockfish single female (Sebastes miniatus). However, many very large fish species (like rougheye and shortraker) are rarely caught when they are gravid in the frigid northern winter waters, which means that the record may be still be broken by a factor of two or more. And, don’t forget that rockfishes are live bearers! Can you imagine having that many children?

Even more amazing is that if a fish reproduces annually for 50 years or more, and has an average of 100,000 per year, she would generate 5 million offspring, none of which ever come back to visit. Why? For a stable population size, only two (on average) out of 5 million will survive to maturity. Oh, the heartbreak! And because there are so many factors involved in their survival, some mothers will contribute several to many adults, but most will contribute absolutely nothing. Again, we are looking at a lottery that has similar odds to our state-run games.

What good are rockfishes?

Their amazing abundance means that as larvae and juveniles rockfishes provide food for larger species. They also make a big dent preying on critters that are smaller than themselves: zooplankton for larvae and juveniles and a potpourri of small fish like herring and sand lance, small crustaceans, squid, essentially whatever appears on the menu. Some adult fish like Pacific ocean perch may continue to prefer a zooplankton diet. From either end of the food web along the eastern Pacific coast, rockfishes are an important component.

So rockfishes are important members of the ecosystem. That makes us feel warm and fuzzy, but what have they done for us lately? Well, they’re really tasty. Many species are caught both by recreational and commercial fishermen (and please note that, in Alaska, females who fish commercially prefer to be called fishermen). The most popular species vary regionally, depending on their abundance.

The commercial fishery for rockfishes began in California in the mid 1800s. San Francisco was the important market and red rockfish were most popular. In the late 1800s, declines in rockfish abundance began to be evident. Before World War II landings in central California averaged about 5 million pounds per year. Following WWII, the fishery expanded rapidly as a result of improved technology; by 1945, landings in California reached 13 million pounds and increased to 18 million in 1958. The fishery expanded north to Oregon and Washington in the 1930s and 1940s, but the fish caught there were used primarily as food for mink. (see http://aquaticcommons.org/3154/1/Diminishing_Returns_Bloeser_OCR.pdf). Recreational and commercial rockfish fishing caused several rockfish populations to decline severely between the 1960s to the 1990s. Some populations have decreased by as much as 98% since 1970 as a result of overfishing and habitat loss. In some areas off southern California, there are very few large adult fishes left on some reefs.

Figure 5. A humongous shortraker rockfish. The 44-inch, 60-pound female shortraker rockfish was caught by the catcher-processor Kodiak Enterprise as it trawled for pollock 2,100 feet below the surface of the Bering Sea, south of the Pribilof Islands. http://www.afsc.noaa.gov/ABL/MESA/mesa_sa_rock_sr.php, and http://www.cbsnews.com/stories/2007/04/06/tech/main2657135.shtml
This decline is well illustrated by Pacific ocean perch (POP). During the 1960s and early 1970s, POP were targeted in an intense foreign (mostly Japanese and Russian) trawl fishery. The harvest peaked in 1965, when it reached 350,000 metric tons (t). Overfishing of U.S. fisheries by foreign vessels led to legislation of the Fishery Conservation and Management Act in 1976 and the subsequent formation of the North Pacific Fishery Management Council (NPFMC). The NPFC defined the abundance of POP as depleted, and in 1978 only 8,000 t were caught. During the period of intensive harvesting, POP stocks were reduced by more than 80% from the virgin biomass throughout their Alaska, Canada, and northern Washington range. A domestic fishery replaced the foreign fleets in 1985, and from 1991 to 1996 the POP fishery was restructured and management practices were changed to encourage the restoration of the POP stock. Since 1996, catches of POP have increased and there is evidence that POP abundance is increasing.

Unfortunately, during the 1980s the West Coast groundfish fishery expanded from a relatively small fishery harvesting sustainable levels to a fishery that was overcapitalized and unsustainable. Many rockfish stocks along the Pacific Coast of California, Oregon, and Washington remain dangerously low even though emergency fishery closures were put in place in 2002.

Are rockfishes particularly susceptible to conservation issues, and if so why? Yes, rockfish populations are particularly sensitive. Many rockfish species have very long life expectancies. Their life history strategy is to produce enormous numbers of offspring during their lives. The idea is that the vast majority of offspring will perish, but that on average, two fish will replace each female to maintain a stable population size. We will examine the basics of production in Chapter 12 (The Importance of Time and Place in Fisheries Management). For now we can look at it from this perspective: if females need to live (for example) 50 years on average to produce two successful offspring, reducing their average life spans by catching large numbers will reduce their ability to maintain the population. There is clearly more to it than that, but let’s wait for Chapter 12.

Another problem is that there are many species. In Japan (the primary market for red rockfish), the primary criteria of value are the red color and size. Although Japanese consumers differentiate among red rockfish species, Alaska fishermen do not because it is difficult to correctly identify every fish in large commercial catches onboard the fishing vessel. So why can’t we demand better enumeration by regulation? The same reason that we have not been monitoring rockfish abundances species-by-species over the last century: money. The long green. Dead presidents. One fisheries manager explained it to me this way: the captain of a fishing boat was asked why we do not have a better handle on the numbers of each rockfish species that is caught. His response was that the market to which he sold only considered whether a rockfish was red or not, so they counted red rockfishes and other rockfishes. The skipper went on to say that there were too many fish and it would cost too much to be more accurate, especially since several of the species were very similar and the observers probably couldn’t tell them apart anyway. He added that if a vessel were told it had to provide those data, they would probably make something up that seemed reasonable.

Is the conservation issue limited to the Pacific Coast of North America? Unfortunately, no. When A.J. Gharrett was in Japan to learn about Asian rockfishes and obtain samples for genetic analysis, he observed two things. First, a number of species were very similar and difficult to identify. Second, time and again his host said that he remembered when this or that species had been much more abundant.
Rockfishes worldwide are experiencing challenges to their perpetuation.

**Summary**

More than 100 species of rockfish occur worldwide. The majority of species are along the Pacific seaboard of North America, although it seems likely that they first appeared in the Asian North Pacific Ocean. Although the reason for the large number of species is unclear, rockfishes are characterized by internal fertilization, which may provide the mechanism for limiting mating possibilities and choices relative to broadcast spawning species. Rockfishes have diverged to occupy numerous ecological niches; and the speciation process continues today, as evidenced by recent separations of cryptic species. Finally, many rockfishes are long lived. But because their offspring sustain high mortalities as larvae and juveniles, rockfish are very sensitive to overfishing. Consequently, rockfishes are a general conservation concern throughout much of their range.
Fisheries managers make the best decisions possible, often with very scanty data. Around every decision corner may lurk a surprise that can alter how they should consider their information. An assumption that is frequently safe is that we can distinguish among species from differences in their appearance. However, when two species diverge from a common ancestor, they may appear very similar until obvious physical differences have been generated by evolutionary pressures. Recently, we discovered that the rougheye rockfish, which has been targeted for harvest in the North Pacific Ocean for at least a century, actually includes two genetically distinct species. This discovery means that management of rougheye rockfish must be revisited and emphasizes that more effort needs to be devoted to understanding the biology of marine species so we can be sure that our conservation and management efforts are effective. In this chapter, we describe the results that revealed the two species as well as subsequent work we have done to learn about the biological differences between the species and how they are distributed geographically.
While we were examining the population structure of rougheye rockfish (*Sebastes aleutianus*) in Alaska waters, we observed some puzzling data. We saw several microsatellite loci (see Chapter 4. Molecular Tools for Population Genetics) that had many more homozygous individuals than would be expected if rougheye rockfish mated more or less at random. This is equivalent to seeing white and red carnations, but no pink carnations (a heterozygous individual that carries an allele for white color and an allele for red color is pink—see Chapter 1. Even Fish Obey Mendel’s Laws). If both red and white carnations interbreed, there should be an abundance of pink carnations.

In many instances, the disparities that we saw in rougheye rockfish occurred in samples of fish caught in the same tow of a trawl or the same longline set. The genetic results we obtained are what would be expected in a mixture of populations (like salmon harvested away from their natal streams) or if more than one species is present. The frequencies of the genotype that we observed at several loci did not follow the Hardy-Weinberg predictions that we considered previously (Chapter 8. How Is Genetics Used for Stock Identification?). One microsatellite locus in particular was striking because its two very abundant alleles were almost always observed as homozygotes. In fact, the results were so highly unlikely that, if you started with that information, you could convert one penny into more than a million dollars with just six spins of the roulette wheel. The short story is that two distinct species have been hiding under a single species name.

There are more than 100 species of rockfishes (genus *Sebastes*) in the world, and more than 30 are common in Alaska waters (see Chapter 9. The Diversity and Challenges of Rockfishes, and Chapter 10. The Wonderful and Diverse World of Rockfishes). Many of these species contribute to commercial or recreational fisheries and all are important components of the marine food web. As part of our research to learn about population structure of commercially important marine species, we are conducting genetics surveys to obtain information about the population structure of rockfishes from Alaska waters. Most of our work is done in collaboration with scientists at the National Marine Fisheries Service (NMFS) Auke Bay Laboratory. We have examined the shortraker and rougheye rockfish and are currently looking at the northern rockfish and Pacific ocean perch (POP). We will describe some of our interesting POP results in the next chapter. Much of our earlier rockfish work focused on mitochondrial DNA (mtDNA), which is passed to offspring only by their mothers, unlike the genes in the nucleus that come from both mothers and fathers (see sidebar 1). More recently we have been studying the variability of diploid nuclear loci, especially microsatellite loci (see Chapter 4. Molecular Tools for Population Genetics). The important point is that mitochondrial and nuclear variations are not merely in independent sets of genes—they are inherited through two entirely different mechanisms. In addition, mtDNA molecules do not exchange genetically (recombine) with other mtDNA molecules. As a result, each mtDNA molecule carries its historic mutational record, which can be used to construct molecular genealogies or gene trees—phylogenetic trees for genes (see sidebar 2). Concordance between anomalous observations of data from the two sources of inheritance is a very powerful signal.

In our mtDNA studies of rougheye rockfish, we saw two distinct clusters of molecular lineages (Figure 1). Each circle in the gene tree represents one molecular sequence (called a haplotype) that incorporates information from all the different restriction sites (simple nucleotide differences—you can review how restriction enzymes work in Chapter 4) that were surveyed. A haplotype is equivalent to a complex allele, but each individual carries only one haplotype because mtDNA is haploid (one genetic copy type per individual). The line segment that separates each circle or dot represents a single mutational difference. Note that 10 haplotypes differ by just a single mutation from haplotype A, but only 5 surround B. You can see in Figure 1 that one large cluster of haplotypes is centered on molecular lineage A and a second cluster is centered on molecular lineage B. The relatively compact size of the lineage B cluster suggests that the individuals, which belong to that family of mtDNA, only recently began to expand after a severe decline in numbers. A population crash is often referred to as a bottleneck and can result in loss of genetic variation. The more severe
the crash and the longer crashes persist, the smaller the amount of genetic variation that can be retained in the population. The larger, more complicated configuration of the lineage A group is characteristic of an older group that has expanded and diverged several times.

We also observed concordant variation at microsatellite loci. The particular alleles carried by a fish at one locus, Sma6, tell us immediately which species we are looking at. Remember that a locus is just the location of a DNA sequence for a particular trait on a particular chromosome; that is, it is a marker. Many such markers have nothing to do with expression of species characteristics (phenotype). They are along for the ride, but they may have been “painted by the species’ brush,” like the trim on a house. Loci that have alleles that can be used to delineate species are referred to as diagnostic loci (Figure 2), because the alleles all by themselves are usually sufficient to tell us (diagnose) which species we are looking at. The Sma6*183/183 homozygote (two copies of the 183 base pair allele) was nearly always seen in fish that had mitochondrial lineages from cluster B, and Sma6*177/177 homozygotes (two copies of the 177 base pair allele) were in fish that had mitochondrial lineages from cluster A. It is extraordinary that virtually no Sma6*177/183 (one each of the 183 and 177 alleles) heterozygotes were observed. In a randomly mating population, one would expect substantial numbers of heterozygotes (see Hardy-Weinberg equilibrium in Chapter 8. How Is Genetics Used for Stock Identification?).
Figure 2. Distributions of allele frequencies for blackspotted (Type 1) and rougheye (Type II) rockfish at eight microsatellite loci. The frequencies reflect the total observed for 225 rougheye (blue bars) and 467 blackspotted (green bars) rockfish analyzed from collections ranging from Oregon to the Aleutian Islands.

Failure to see $^{*}177/183$ heterozygotes indicates that rougheye rockfish that possessed the Sma6$^{*}177/177$ genotype did not mate with individuals that had the Sma6$^{*}183/183$ genotype, even though both were present in many areas.

Let’s consider what happens when two species newly separate. Because they diverged from a single common ancestor, they have nearly identical instruction manuals (DNA sequences). Divergence often occurs if they are separated in time, geographically, or behaviorally; and mechanisms that isolate them reproductively and genetically develop.
At the time of divergence, they may be very similar genetically and morphologically like the species flock of cichlids in Lake Malawi (see Chapter 9. The Diversity and Challenges of Rockfishes). For a long time after two new species have diverged from their common ancestor, they continue to carry many of the same alleles (DNA sequences) for genes. However, as a result of genetic isolation, finite population sizes (which leads to random divergence), and accrual of mutations, the two species will eventually diverge genetically until they carry distinct alleles for most of their genes, although the underlying DNA sequence motif will still exist even between species as diverse as amphioxus and humans (see Chapter 9). The process during which their allelic complements become dissimilar is referred to as lineage sorting (see sidebar 2). In this usage, we are describing the molecular lineages of the alleles themselves. Other microsatellite loci show various levels of divergence, which indicates that lineage sorting is incomplete at those loci (Figure 2); but their differences also corroborate the story presented by SmS6. Both types of rougheye rockfish occur in many of the same areas in the northeastern Gulf of Alaska and are often caught in the same trawl hauls or longline sets. The microsatellite evidence documents genetic isolation, and the mitochondrial data document the isolation and that the two forms have been isolated for an extended period. Distinct species that are difficult to tell apart are often referred to as cryptic species. We were not the first to suggest that there is more than one species of rougheye rockfish; the existence of cryptic species of rougheye rockfish was suggested by other genetics studies based on allozymes (see Chapter 4. Molecular Tools for Population Genetics), which are nuclear loci. What our study contributed was the strong divergence demonstrated by the mtDNA data that correlated with the nuclear microsatellite differences. Dr. J. Orr has assigned Type I fish the name blackspotted rockfish (Sebastes melanostictus) and kept the name rougheye rockfish for Type II fish (S. aleutianus).

Why have these species been difficult to distinguish? They are obviously very similar looking. We can distinguish them by using genetic markers, but genetic markers are not always convenient to apply, especially under field conditions. The challenge is to identify morphological characteristics that can be used to distinguish between species when they are caught, rather than later in a lab. So, we looked to see if we could find characteristics that distinguish the two types. It had been reported that there might be coloration differences. Indeed, rougheye rockfish coloration varies from very pale with little or no mottling to nearly black. Fish caught along the Aleutian Islands are very darkly colored and almost exclusively blackspotted rockfish. Fish caught in the northeastern Gulf of Alaska include blackspotted, which range from dark to reddish with blotches to very light, and rougheye rockfish. We conducted a study that examined the coloration of fish sampled from sites near the Queen Charlotte Islands, in Southeast Alaska inside waters, in Prince William Sound and Kachemak Bay, as well as from the outside waters of the northeastern Gulf of Alaska and along the Aleutian Islands. The conclusion was that all rougheye rockfish are very lightly colored and most, but not all, of the blackspotted are dark or reddish colored or have some pigmentation blotches (Figure 3). More recent studies confirm that a not-insubstantial portion of the light-colored fish are blackspotted rockfish. In our experience, coloration provides a good first cut at identifying the species, but
What Does Genetics Have to Do with It?

Sidebar 1

MITOCHONDRIA

Mitochondria are small subcellular structures that are responsible for aerobic (oxygen) metabolism in higher organisms (fungi, plants, animals) (Figure 4). The primordial atmosphere had little oxygen and the first organisms were anaerobic, like septic tank microbes. In fact, oxygen was toxic to most of those organisms. Some microbes developed the means to detoxify oxygen by converting it to carbon dioxide and water. Mitochondria probably descended from these early aerobic pioneers, and were later adopted by early complex cells for their ability to detoxify oxygen. Subsequently, mitochondria evolved into structures that used oxygen in generating biotic energy, which made aerobic metabolism possible—aerobic metabolism is much more efficient than (primitive) anaerobic metabolism. Mitochondria have been enslaved ever since. The mitochondrial genome carries a small number of genes in a circle of DNA (Figure 5). Those genes are necessary for the aerobic function of the mitochondrion; they and the apparatus that is needed to express those genes as proteins probably descended from the ancient microbial ancestor.

Figure 4. Electron micrographs of mitochondria (top) and mitochondrial DNA (bottom). mtDNA photo courtesy of W. Brown.

Figure 5. Circular gene map of vertebrate mitochondrial DNA. The letters inside show locations of the 22 transfer RNAs that specify amino acids in protein synthesis. On the outside are genes that specify proteins (ND1, ND2, ND3, ND4, ND4L, ND5, COI, COII, COIII, and CytB), genes that specify ribosomal RNAs that are part of ribosome structures involved in protein synthesis (L rRNA and S rRNA), and the control region, where DNA synthesis is initiated in replicating the mitochondrial genome.
Sidebar 2

THE BEAUTY OF MITOCHONDRIAL DNA IN PHYLOGENETIC ANALYSES

Most of the genes needed to construct and operate mitochondria are carried in the nucleus of the cell that has the mitochondrion. Those diploid genes are passed on through the normal meiotic process (see Chapter 1. Even Fish Obey Mendel’s Laws). The few genes carried by the mitochondrion, however, are passed directly from mothers through the egg cytoplasm to offspring. In contrast, paternal mtDNA is almost never passed to offspring and in some species there may be mechanisms in the egg that destroy paternal mtDNA. Consequently, the offspring receive from their mother complete, intact copies of mtDNA, which have no opportunity to exchange genetically with other mtDNA molecules. The only changes that occur are the occasional mutation. If the mutation is not deleterious or favorable, it will be passed to offspring.

Each mutation creates a new lineage that is distinct from every other lineage. Recurrent mutations (independent mutations at the same site in many different individuals) are rare because there are usually more than 16,000 nucleotides in mtDNA genomes, and a single nucleotide change can go from a single nucleotide to any of three other nucleotides (for example, from adenine to thymine, cytosine, or guanine). A recurrent mutation would be more rare than lightning striking twice in the same spot; golfing legend Lee Trevino is the rare individual who has been struck by lightning not once, but twice! Note that as a result of a single nucleotide change (mutation) a mtDNA molecule differs from its predecessor by only a single nucleotide. As mutations accrue over time in a lineage, the collection of lineages in a population carry more and more mutational differences—they diverge over time. The number of mtDNA lineages that can exist in a population is constrained by the number of females in that population, because only females can pass on their mitochondria. In fact, each time a population crash occurs, a new (lower) maximum number of lineages is set. The idea is that as new lineages are being created, others are being lost at random. The random loss of lineages that occurs because of constraints in the number that can be accommodated in a population is referred to as lineage sorting (Figure 6).

A gene tree can be constructed from the lineages. Two assumptions are required: (1) mutations accrue one at a time; and (2) the simplest (most parsimonious) relationships between lineages are the most likely. Here is an example of parsimony: Georgie didn’t turn in his homework. Which is more likely (parsimonious) explanation? (1) He did not do it; or (2) his explanation that a creature from outer space swooped in and stole it because it was such a brilliant piece of work.

In some instances, an unobserved lineage must be inserted into the tree to maintain those assumptions and connections between the lineages that were observed. The missing types would be analogous to the gaps in the fossil record (see Chapter 9. The Diversity and Challenges of Rockfishes). For example, in Figure 6, lineages 2, 3, 4, and 5 are all observed in the most recent generation (surrounded by the ellipse). From their mutational differences, it can be deduced that some appeared more recently than others—lineage 2 is the ancestor of 5 and lineage 4 is the ancestor of 5. To bring the two branches together in the gene tree, however, we must presume that there was an earlier unseen type—lineage 1. Note that information of the very first lineage has been lost, and we can only go back as far as our information takes us. Also note that the gene tree itself was constructed in a manner that the most recent lineages are at the tips of the tree and the oldest (lineage 1 in this example) is in the interior. Going from the tip to the interior, we go back in time, although the time scale is not necessarily linear.

We have just considered three different ideas that can be used to interpret mtDNA variation. The first is that recently divergent lineages differ by very few mutations. The reverse of that idea is that lineages that separated a long time ago would be expected to have many mutational differences. There is an element of time in the amount of divergence that is observed between two mtDNA molecules. The second idea is that after a population crash, the number of lineages will be reduced. Then, as the population expands again, more lineages can be accommodated, but these will emerge as recent mutations, so there may be a cluster of very similar lineages in the genealogy. Finally, as a result of lineage sorting over a long time, two lineages will diverge and one will eventually disappear. A consequence is that even if genes (such as mtDNA) did not contribute to the speciation event, after two sibling species have been
isolated for some time they will be characterized by different mtDNA lineages.

Gene trees carry information about the history (recent geological time) of populations. Different demographic histories generate different patterns. We will examine gene tree patterns that are characteristic of four different events (Figure 7). Pattern A: a population crash or bottleneck severely reduces the number of lineages that can be maintained in a population. If the low population size persists, other lineages may be lost because there cannot be more lineages than there are individuals, and each generation is a lottery for which lineages will be carried forth by random chance. This is another aspect of random drift that we considered earlier (see Chapter 2. How Genes Vary in Populations). The lineages that disappear are losers in the genetics lottery, purely by random chance.

However, when the population begins to expand to larger sizes more lineages can be carried. Of course the source of these is recent mutation and all the new lineages will be very similar to the type(s) that survived the bottleneck. The pattern of 10 haplotypes surrounding haplotype A is often referred to as a star phylogeny. Pattern B: a stable population can only maintain so many lineages. However, there is no time constraint on the relationships among the lineages and as time goes on, star phylogenies degenerate and the pattern of relationships becomes less distinct and many "missing" lineages must be inserted. Pattern C: geologic, oceanographic, or demographic events can break populations into isolated components. If the isolation persists, the older lineages may show the historic relationship, but new lineages will be population specific. Eventually, the isolation will result in complete lineage sorting and the populations will differ entirely. Note that complete lineage sorting may not even be attained for some time in sibling species. Pattern D: two populations that are largely, but incompletely, isolated will show divergence, but usually a quantitative divergence in the abundances of the different lineages that they carry. As long as there is some exchange, however, the strong divergence illustrated by Pattern C will not occur.

**Sidebar 2**

Figure 6. An illustration of lineage sorting, the true genealogy, and a gene tree based on detectable lineages. Mutation events generate new lineages, but constraints in the numbers of females in the population results in random loss of some lineages. The connected circles at the lower right are the typical way in which the information from the detectable part of the genealogy is presented. This is a very simple tree because it is not possible to realistically depict all of the lineages that make up most populations.
Figure 7. Examples of gene trees showing the results of different demographic processes. Trees in A and B represent haplotypes within a single population resulting from two different demographic histories. Trees in C and D reflect distributions of haplotypes between two populations resulting from different historic and current processes. For example, the blue population in C has individuals of all the haplotypes in the left side of the tree and of the central haplotype in the right side. The green population, in contrast, has individuals of only the central haplotype on the right, but of all the haplotypes on the right side of the tree. Both populations have both haplotypes in tree D, but the blue population has a higher abundance of haplotypes on the left side and the green population has a higher abundance of the right side.
it is not good for 100% identification.

The samples that we studied are referred to as “samples of opportunity,” which means that the samples were the byproduct of periodic stock assessment surveys conducted by the National Oceanic and Atmospheric Administration (NOAA) and from other surveys conducted by the Alaska Department of Fish and Game and the Department of Fisheries and Oceans Canada. The designs of these surveys did not consider the questions that might pertain directly to the rougheye rockfish species. Many different projects are usually piggybacked on activities like these surveys, which is good for the taxpayer; and often the sampling design does not affect the study. NOAA periodically conducts two separate stock assessment surveys, one by trawl and the other by longline. Trawl surveys tend to be conducted in somewhat shallower water that has a fairly regular bottom, because large obstacles like boulders can foul and damage the gear. Longlines can be set along very irregular bottoms and at greater depths. The samples that we examined reflect the geographic distributions of the two species; from their incidence in the catches, we can roughly describe their distributions (Figure 8). Clearly, blackspotted rockfish predominate in the collections sampled west of Kodiak and along the Aleutian Islands. Both types were collected in the northeastern Gulf of Alaska, which appears to be the center of rougheye rockfish abundance. Also, only blackspotted were observed at Surveyor Seamount, one of a string of pinnacles (an underwater mountain range) that are out in the middle of the Gulf of Alaska. In other areas, it is not clear which form predominates. The reason is that in many areas only longlines or trawls were used, not both. Because they probably sample different depths and habitats, they would not necessarily sample the two species evenly if they were distributed in different depths or habitats.

It is unlikely that two species, which inhabit overlapping ranges, will occupy exactly the same ecological niche. They (or their young) will most likely exploit habitats at different depths or in different locations, or partition a shared habitat in different ways. That is, they will have different life histories, which may include food habits or preferences. As a result, there may be differences in their morphology that would relate to their roles in the ecosystem. That perspective provided us with additional avenues to explore for differences between these cryptic species. So, in addition to examining coloration, we (K. and T. Mecklenburg) took morphological measurements from fish that were caught in the northeastern Gulf of Alaska and that had been identified with mtDNA and microsatellite markers. That study showed some consistent differences in the averages of the two types. Unfortunately, the ranges of the counts and measurements for the two species broadly overlapped. The result is that no single count or measurement that we examined could be used to tell them apart. However, the differences in averages indicated that rougheye tend to have more numerous and longer gill rakers than blackspotted. Gill rakers are small bony processes that project inward toward the mouth and throat from the gill arches and are used to sieve the water for very small organisms or to clamp down on larger prey that are being swallowed to prevent them from backing out. Clearly, the length and robustness of gill rakers may be involved in prey selectivity. For example, planktivorous sockeye salmon generally have longer and more numerous gill rakers than other Pacific salmon species, which feed on squid or small fish like herring or sand lance. In addition, the rougheye rockfish tended to have deeper bodies (several different coordinated measurements) and shorter dorsal spines than the blackspotted rockfish. Both the gill raker and shape differences suggest that the two species are ecologically divergent as well as genetically divergent.

We also took a preliminary look at their habitat preferences. It was preliminary because, as pointed out above, the samples we had to work with were samples of opportunity. For our samples, the sampling design for the assessment surveys may have influenced the apparent distributions of the rougheye rockfish species. There are two important questions that we would like to ask: Do the two gear types have different catching efficiencies for the two species? and Are the two species segregated by depth and or bottom type? Our analyses suggested that gear and depth may have been related to the distributions of species caught, but without parallel sampling with both types of gear, it is difficult to conclude unequivocally that the depth and/or bottom type
separate the two species. We did notice, however, that although both types of fish were caught in many hauls, in most of the large hauls (at least 10 fish were landed) longlines caught more blackspotted rockfish and trawls caught more rougheye rockfish, which strongly suggests they segregate to some extent by bottom type or depth, unless the two gear types have substantially different catching efficiencies (Table 1). Of course, experiments designed to address these questions must be conducted in order to arrive at firm conclusions.

Our genetic observations clearly demonstrated that there are two species, and often both are caught in the same places, especially in the northeastern Gulf of Alaska. Since our studies, there have been additional efforts to identify morphological cues to distinguish the two species and an effort to evaluate the accuracy of those identifications by comparing them to genetic identifications is ongoing. In addition, because the laboratory analyses are a bit tedious, we have developed SNP markers (see Chapter

Figure 8. Numbers of blackspotted (top) and rougheye (bottom) rockfish collected in the accumulated surveys by the National Oceanic and Atmospheric Administration, Alaska Department of Fish and Game, and Department of Fisheries and Oceans Canada. Species were identified by combined mitochondrial and nuclear genetic markers. The ellipses represent the geographic areas covered by the hauls that were combined to produce the numbers shown. In some hauls, both types of rougheye rockfish were caught.
What Does Genetics Have to Do with It?

Table 1. Distribution of rougheye and blackspotted rockfish in longline and trawl catches east of 155°W that had at least 10 fish total. Hauls made west of 155°W were not considered because very few rougheye rockfish were captured in those hauls.

<table>
<thead>
<tr>
<th>Proportion</th>
<th>Longline</th>
<th>Trawl</th>
</tr>
</thead>
<tbody>
<tr>
<td>All blackspotted</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>≥0.75 blackspotted</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>More than 0.25 and less than 0.75 blackspotted</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>≤0.25 blackspotted</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>All rougheye</td>
<td>1</td>
<td>11</td>
</tr>
</tbody>
</table>

4. Molecular Tools for Population Genetics) to replace the diagnostic mitochondrial markers for distinguishing between the two species. With those new tools, we will be able to analyze many individuals in just a little more time than it takes to isolate the DNA.

The other interesting question is: How do the two species differ in their ecologies and life histories? Because of the disparity in the information that might be obtained from trawl and longline surveys, one of the approaches will be to sample in parallel with both sets of gear. Unfortunately, the experiments can be conducted only in areas that can be sampled by both types of gear, which limits sampling to less rocky, shallower areas that have a relatively low gradient. However, if the catches are similar, it might be presumed that there are no strong gear-related differences. Otherwise, it will be more challenging to evaluate longline and trawl data. Another approach is to make submarine observations in areas that appear to have large concentrations of one species but not the other, and observe them in their habitats to see if there are obvious habitat-related behavioral differences such as schooling or association with different substrates.

These and other approaches will help us learn more about adults, but there may be important differences between their larvae and juveniles, and we do not yet have effective methods to track rockfishes during their early life histories, when most of mortality occurs.

Summary

The concordance of diagnostic markers that included both nuclear loci and mtDNA revealed that there are actually two different species hiding under the name of rougheye rockfish. They have now been separated into blackspotted and rougheye rockfish. This genetic evidence is especially compelling, because nuclear and mitochondrial markers are inherited through two different mechanisms. Because a rapid means to identify the species on board the survey vessels would simplify surveys of these species, we compared them morphologically. Although most rougheye rockfish are very lightly pigmented and have no speckling or mottling, some of the blackspotted rockfish also are very light with no pigmentation, although most have a darker coloration. We also observed that the two species differ on average in numbers and sizes of gill rakers in body shape and the ranges of the counts and measurements overlap; we did not find a character (or small set of characters) that was 100% effective for identifying species. The morphological differences do indicate, however, that the two species may occupy different ecological niches. Examination of the catch data suggests that longline sets have more blackspotted and trawl hauls have more rougheye rockfish. Trawls are deployed in shallower water with fewer obstructions than longlines; longlines are often set at greater depths and can be used in very rocky areas. These differences suggest that the two species may occupy different habitats, but there have been no comparisons of the relative efficiencies of the gear types in a common habitat. By far the highest mortality occurs during the early life of most marine species. Unfortunately, as yet we have no way to track larvae and juveniles of many species during this critical part of their life histories.
Up to now, we’ve looked at genetics and fishes in a number of ways. We’ve examined
the principles of the science of genetics (Chapter 1. Even Fish Obey Mendel’s Laws);
looked at the forces that generate genetic changes (Chapter 2. How Genes Vary
in Fish Populations, Chapter 3. History of a Salmon Population, and Chapter 5.
Fish Population du Jour); explored the tools that are used by geneticists (Chapter
4. Molecular Tools for Population Genetics); and examined several applications
of genetics (Chapter 6. Genetics and Hatcheries, Chapter 7. The Lowdown on
Frankenfish, Chapter 8. How Is Genetics Used for Stock Identification, and
Chapter 11. Is a Rougheye Rockfish Really a Rougheye Rockfish?). In this chapter,
we examine the physical nature of the Gulf of Alaska. The short term and long
term success of species depends on their ability to adapt to and flourish in this
complicated environment.

Long term (evolutionary time) persistence of a
species ultimately depends on the ability of multiple
populations to succeed, which means that they must
be able to adapt to changes in their environments. In
the short term (seconds to hours or days), they must
be genetically “wired” to respond physiologically
to cues in their immediate environment, such as
presence of predators or prey, or to physical changes
such as chemistry or temperature. The genetic
“wiring” both designs and maintains their nervous
and hormonal systems. For example, adrenalin is
one of the many hormones that are always present
just waiting for something to do. In the longer
term, genetic changes that make a population more
productive can result in local adaptation (see Chapter 5). And in the very long term, environmental changes may occur that result in the emergence of new species (e.g., Chapter 11).

We have talked fairly loosely about the "environment," but we are long overdue in paying attention to an environment in more detail. Obviously, the interaction of an individual and a population with their environment is key to their success (producing more of their kind) or failure (extinction); and each species, population, and individual has unique environmental experiences during their lifetimes. There are, however, some environments that are common to many species. What we will learn is that environments can be very complicated, even if we ignore year-to-year variation.

In this chapter, we examine the Gulf of Alaska (GOA), which is the environment for a wide variety of species such as rockfishes and (for substantial parts of their lives) salmon. If the GOA were a big bathtub in which the water just sat, it would be easy to describe. However, it is more like a very complicated Jacuzzi that has many currents and is surrounded by and filled with many mountains as well; each of these features influences the environment. Also, multiple sources of water feed the GOA and it drains into other bodies of water such as the Bering Sea and northern Pacific Ocean. The study of oceans is called oceanography and includes its biological, chemical, and physical properties. My knowledge of oceanography is limited, especially physical oceanography, and it is about time I learned a bit more. So as I learn, I will take you on a personally guided voyage that focuses on the GOA. The good news is my coauthor, Dr. Tom Weingartner, is a physical oceanographer: a specialist in understanding the physical factors that describe and drive ocean systems. Tom will ensure that what we learn is correct and current.

So where do we start? The first thing we must do is learn what physical factors and processes characterize oceans. In particular, we are interested in currents, tides, physical habitats, and sources of nutrition. What we will find is that these factors cause ripples (a little joke) that connect all of the earth's oceans and create large subareas, such as the GOA, that can be looked at as individual entities, although they are all connected to some degree. Next, we'll look at the common connecting processes and finally we'll examine the GOA in more detail.

**Physical factors responsible for oceanographic features**

The earth's surface is about 30% land, which is unevenly distributed around the earth and scarred with features such as mountain ranges, canyons, and deserts. The remaining 70% is water, mostly seawater. The earth is surrounded by an atmosphere, which is primarily nitrogen and oxygen but also has important amounts of carbon dioxide (CO2) and gaseous water (H2O). We will see that atmospheric and oceanic processes are very tightly linked. Three factors exert enormous effects on the atmosphere and oceans: the sun, the earth's primary energy source; the earth itself, which has (in a human life time scale) a relatively fixed shape and spins on its axis; and the moon.

**Effects of the earth**

The earth itself has multiple effects on the ocean. First, the earth rotates around its axis; looking down from above the North Pole, it rotates counterclockwise. As we stand on earth and look outward, we are aware that night follows day follows night follows... If we look at the earth from space, we see that even while it rapidly circles (orbits) the sun, it rotates around its own axis, spinning like a top. The fattest place on earth, the equator, moves the fastest—almost twice as fast as commercial jet airplanes fly. So if you want to fly west, why can’t you just take off and hover and let the earth move beneath you? The answer is that because of friction the rotating earth carries the atmosphere with it. If it did not, we would experience colossal winds all the time, except at the North and South Poles. In fact, there are prevailing winds that blow toward the east, so it actually takes more time to fly west than east because of these “headwinds.” The rotation of the earth does have an important phenomenon called the Coriolis Effect, which is related to the different rotational speeds of the surface of the earth at different latitudes (see sidebar 1). Another important effect is that the axis about which the earth turns
(North Pole to South Pole) is not straight up and
down relative to the sun, but tilted at about 23.5°. That
tilt is responsible for the seasons: during summer in
the Northern Hemisphere, the North Pole is tipped
toward the sun and during the winter it is tipped away
(sideways).

Other processes also affect the ocean. As we are
aware, the surface of the earth is very irregular. The
land is characterized by mountain ranges, valleys,
plains, and plateaus. Terrestrial elevations dip to 1,371
feet below sea level along the Dead Sea coast and
rise to 29,029 feet (nearly 5.5 miles) at the top of Mt.
Everest. The seabed is also irregular and characterized
by the same kinds of features. The deepest point, the
Challenger Deep in the Mariana Trench, plunges to
35,797 feet, more than 6¾ miles deep. Although the
shapes are ever changing, the time scale for change
is too long to be significant for most day-to-day or
year-to-year processes. However, the changes clearly
influenced the evolution and distributions of species
as we see them today (see Chapter 3. History of
a Salmon Population). Finally, the sun and moon
influence both water and atmospheric movements on
earth.

Effects of the sun

The sun is essential for life on earth. Without the sun,
the earth would be a frigid chunk of rock floating in
the universe. Energy from the sun fuels most living
things, either directly for many plants or indirectly
for things that eat plants or animals that eat things
that eat plants. The sun warms the earth, so that water
over much of the earth is liquid: streams, lakes, rivers,
and oceans; rain; and about 70% of our bodies. In
some areas of the earth or during certain times, water
is solid in the form of ice and snow. Water as a gas
is important for cycling water between aquatic and
terrestrial habitats. In this chapter, we will see how the
sun produces these and other oceanographic effects.
The sun is important for influencing the weather as
well as the oceans, and the effects of weather and
oceanography cannot always be separated.

Let’s start with a simple, but important, example.
You are on a boat at the equator, say in the middle
of the Pacific Ocean. What is it like? Generally hot.
Why? Because even with seasonal variation, the sun
shines straight down at midday. The atmosphere is
relatively transparent to the sun, but when light hits
the water from directly overhead, most of its energy is
absorbed rather than reflected by the water’s surface,
and converted to thermal energy (heat). Because
you’re in the middle of the ocean, sun-warmed water
is all around you. The warming water also heats the
air at the water’s surface, and causes some evaporation
(changing water from liquid to gas). What happens
when air is warmed? Just as with the colorful hot air
balloons (for example in the wine region of Napa
Valley, which take tourists up in the air to enjoy the
view) heating air causes it to expand so it is less dense
(lighter) than the cooler air around it. The more
buoyant balloon rises into the sky. Because air thins
(becomes less dense) as you ascend until there is
none at all, the balloon will eventually climb to air of
the same density.

Similarly, the warmed air at the equator will
ascend until it reaches air of the same density. The
newly arrived air displaces air already there and
pushes it north and south. At the higher altitude the
warmer air cools, but it moves from the equator (0°
latitude) to about 30° north or south latitude until
it has cooled sufficiently to become heavy enough
denser) to sink closer to the earth, where it is again
warmed as it descends. The air at the equator is
replaced along the sea surface by air from the north
and south. The result is a circle of air that produces a
net wind at the equator. But because of the Coriolis
Effect (sideways), the winds blow regularly from east
to west near the equator. These “trade winds” (see
sideways) push water before them to produce the
North Equatorial Current. If you don’t believe that
wind can drive currents, fill up your washbasin and
blow on the surface along one of the sides. You can
even use drops of food coloring to help you track the
motion. After a little huffing and puffing, you will
see that the water is moving around the bowl in the
direction you are blowing. Notice also that waves
are created by the friction between the air and water.
Bigger winds produce larger waves. I use the size
and number of whitecaps (frothy tips of waves) that
form in a brisk wind to help me decide if I want to go
fishing in my small boat.

The cold climates in the far north and south cool
THE CORIOLIS AND OTHER EARTHLY EFFECTS

Let’s start this sidebar with a practical example of what “frame of reference” is and how objects that rotate can influence perception. You are the target of an assassination attempt. The assassin is following you, waiting for an opportunity. He only takes headshots. As luck would have it, you are near an amusement park that has one of the world’s fastest carousels (18 mph—10.3 revolutions per minute counterclockwise. Hang on!). You get on and the assassin mounts the second horse behind you and begins to take aim with his 38 special, which has a muzzle velocity of 600 feet per second (fps) (Figure 1). To avoid the attention that would be generated by your falling off a stationary horse, he waits until the carousel is rotating. You remain calm as he takes dead aim because your high school trigonometry class enabled you to quickly calculate that the speed of your horse rotating around the center of the carousel will move you out of the path of the bullet. In fact, the bullet will miss the center of your noggin by a bit more than 6 inches—you will feel the bullet buzz by your right ear (Figure 1).

So what happened here? Once the bullet is fired, it goes straight in the direction it was aimed with the combined velocity of the carousel and the muzzle velocity. It does not curve! Curving would require a sidewise force on the bullet, which no longer exists once it clears the barrel. If the assassin had fired before the carousel was turning, you would be out of luck. If you were both on bicycles going in a straight line, again you would be toast. The idea is that once something (like a bullet) begins to move, it goes straight. Of course, things like gravity and air resistance will alter its path.

We can use this example to define “frame of reference,” a term that describes how things look from a particular position. Spectators watching the carousel will see you and the assassin rotating around counterclockwise, but to you on the carousel, the spectators are moving clockwise around the carousel. From your (and the assassin’s) frame of reference, when the bullet is fired it starts straight at you at a speed of 600 fps, but appears to curve to your right. Actually, the bullet is going straight and you are moving to the left. To the spectators, the bullet travels straight from the gun at a speed of 600 fps plus the speed of the carousel when the gun was fired.

Another thing we need to know is that on rotating objects, the farther you go from the center the faster they go. If you had been on a horse closer to the center of the carousel, you probably would have been hit because the horse would not have moved fast enough to get you out of the way. Imagine an old-fashioned vinyl record on a turntable. Records rotate clockwise at 33⅓ revolutions per minute. They measure 6 inches from their center to the outside edge. That means that a point at the edge moves about 1.19 mph. A point midway between the center and the edge moves half as far around each revolution, so it moves half as fast. The same idea works with rotating spheres (like the earth). Look at the earth from above the North Pole. It looks like a counterclockwise rotating disc. Although it would seem that latitude 45°N is halfway between the equator (outer edge latitude 0°) and the North Pole (the center at latitude 90°N), points at latitude 60° are actually the midpoint on the
2-dimensional disc seen from above (Figure 2) and move half as fast as points on the equator. If you look from above the South Pole, the earth looks like a disc spinning clockwise.

We know that the earth rotates once per day relative to the sun. If a satellite over the North Pole is the frame of reference, the earth is spinning counterclockwise; that is why the sun rises in the east and sets in the west. As in the example of the vinyl record, a person at the equator has to move a greater distance to complete a rotation than a person at 60°N latitude or one at the North Pole. Movement at the equator is 1,037 mph, at 60°N a relatively sluggish 519 mph, and at one of the poles, no movement—just one spinning in place once a day, like the earth. Of course while it is rotating, the earth itself also moves at 66,600 mph in its orbit around the sun (Figure 3).

The primary components of the earth are the land, waters, and atmosphere. The shape of the land on earth changes in geologic time as a result of continental drift, uprisings and eroding mountains, advance and retreat of glaciers, and so on. In the short term the land has a relatively fixed shape and its features rotate with the earth.

In contrast, both the oceans and atmosphere are fluid, and while they tend to move with the earth’s rotation (or else we would have 1,037 mph winds and currents at the equator), they can also move in other directions. The earth’s rotation produces another, not so obvious, effect on these fluids called the Coriolis Effect. Two factors are involved. The first is inertia: a body tends to maintain its motion unless a force acts on it to change its direction or speed. Moving a fixed object (say a boulder) takes effort (force). An object such as an arrow moves in a straight line, but eventually the force of gravity makes it plunge to the ground.

The north-south axis of the earth is tilted about 23.5° relative to the sun. It maintains that angle throughout its annual circuit, although it changes over geologic time which contributes to the advance and retreat of glaciers (Chapter 3. History of a Salmon Population). During the winter, the Northern Hemisphere is tipped away from the sun and cools; during the summer, it is tipped toward the sun and warms (Figure 4). Of course the changes in season influence atmospheric and oceanographic conditions, especially nearer the poles.

The second factor is that both air and water are (by definition) fluids. An object moving through them moves straight, but may be influenced by frictional forces (drag) and gravity. Now we need to recall the example of what happened in the assassination attempt on you. Because you were astride your steed,
which was on the rotating carousel, the bullet missed you because it travelled straight toward where you were when it was fired. The same thing occurs in air transportation. Let’s fly from Honolulu to Anchorage, which is 4,230 miles—a 5.75 hour flight. We will observe the flight from a satellite hovering over a fixed point on the earth (like a communications satellite). In our frame of reference, the earth will not seem to be spinning, but the airplane will be moving over the earth’s surface. When the plane takes off from Honolulu, it is aimed directly at Anchorage (blue path in Figure 5). If it maintains that direction, it will be near Hudson Bay in Canada when it arrives at 61.3°N (Anchorage’s latitude). The Honolulu airport is spinning at 966 mph, but the more northern Anchorage airport is moving at only 499 mph. When the plane takes off, it has an air speed (485 mph relative to the earth) plus the speed it had by being in Honolulu. Its west to east speed is 467 mph faster than Anchorage. During the 5.75 hour flight to 61.3°N, it would travel an additional 1,730 miles east of Anchorage.

Mathematical modeling of a physical process generally proceeds by considering the factors that are involved in the process one at a time. Ordinarily, the most important factor is considered first (often, analyses are conducted to determine which one is most important) and then other factors are brought into play in order of importance until the model can reliably describe or predict the process. In our model of tracking an airplane from Honolulu to Anchorage, we only considered the very simple trajectory of the plane over the rotating earth. By ignoring the atmosphere, we did not consider things like aerodynamics—both the need for air to fly and the plane’s resistance to flight. The concept that we observed, however, is that the paths of things (planes, water, air, etc.) moving around the earth are influenced by the spinning of the earth itself. Moreover, the path of our plane bent to the right (east) of its intended path. If we had flown from Anchorage to Honolulu, we...
the oceans, often freezing it at the surface. Colder (liquid) water is denser than warmer water and tends to sink. Salt lowers the freezing point of water; and when saltwater freezes, the ice often leaves the salt behind. The result is higher salinity water at or near the freezing point on which the ice floats. The colder, more saline water displaces the warmer, less saline water below it and initiates an underwater current. Note that the influence of the sun at the equator primarily generates surface currents, whereas the effects of cold climates at the poles generally produce subsurface currents.

Effects of the atmosphere—weather
Weather has a profound effect on the oceanography of the Gulf of Alaska. Water has a very high specific heat capacity. Specific heat capacity is a measure of the amount of energy (heat) that is required to increase the temperature of a substance. Liquid water has one of the highest heat capacities of all substances (1 calorie per gram per degree Celsius (cal/g/°C), versus 0.11 cal/g/°C for cast iron, 0.41 for wood, 0.19 for quartz sand, and 0.5 for ice). In addition, it takes energy to evaporate it from liquid water (540 cal/g) or to melt it from ice (80 cal/g). It should be clear that water tends to remain liquid unless a lot of heat is added to or removed from it, and that liquid water does not change temperature very rapidly. That means water is an excellent source of or acceptor of heat (heat sink) from the land or the atmosphere, and once warmed, an excellent heat donor.

So how does the specific heat of water translate to weather? In the far north, winter temperatures cause the land to lose heat (become colder), which in turn cools the air and makes it dense (heavy). Over the sea, however, a different situation occurs: the air is warmed by the ocean (to 32°F, 0°C) and is warm compared to the subzero polar temperatures over the land and ice, so is not as dense. The weight of the air above any particular location is the atmospheric pressure there (measured by barometers). In fact, the barometric pressure can have real effects on the sea surface level. Sea levels are actually higher where there is low barometric pressure and lower where pressure is high, and those levels can be measured remotely by satellite imaging. These pressure changes in sea level are in addition to tidal cycles.

An area of low air pressure, known as the Aleutian Low, often sits just south of the western end of the Aleutian Islands during the winter and influences weather over the entire region. Differences in atmospheric pressure produce the same kind of effect that we see in the Hadley Cell (sidebar 2) because air tends to move from high pressure to low pressure. Air movement around low pressure areas is counterclockwise because the Coriolis Effect causes the air rushing toward the low pressure area to deflect to its right (Figure 13). Consequently, prevailing winds generally blow toward the southwest along the Kamchatka Peninsula and to the northwest along the British Columbia and Southeast Alaska coasts. Those winds produce surface currents in the ocean that flow in the same direction as the wind.

At our house in Juneau, the prevailing winds are from the south, just like the effect of the Aleutian Low would predict. The house, which fronts on a saltwater cove and faces northwest, is protected from those winds. Occasionally in the winter, however, a cold, high pressure system will descend from the north.
Atmospheric conditions are responsible for many oceanographic processes. Of course, energy from the sun drives a lot of them. We'll start at (or near depending on the season) the equator where the sun beams straight down on the water. Sunlight readily penetrates the atmosphere, but water absorbs energy from the light that is not reflected. Consequently, the water at the equator heats up and warms the air immediately above it; some water evaporates into the atmosphere. Warm air is less dense (lighter) than colder air and can carry more water vapor. The warm moist air rises, like a hot air balloon. As it rises, it cools and the water condenses to rain. So near the equator, it is often rainy. In some places, the weather cycles between hot and rainy seasons, a result of the tilt of the earth relative to the sun. But our focus is on the rising warm air. Gravity prevents it from rising too far and as it cools it becomes colder and denser (heavier). However, other air rising below it pushes it out of the way, and because the upper air mass can only rise so far, it is pushed toward the poles. As it moves north (or south) it continues to lose water to rain until it dries thoroughly and, as it cools, becomes heavy enough to sink back toward the earth. The downward movement of dry air takes place at about 30°N (or south) latitude, you will see many deserts: in Africa (the Sahara and Kalahari), Australia, Mexico, India, and Chile (Atacama Desert). Why didn't I learn about this from junior high science or geography?! It's really cool stuff!

The second result of the Hadley Cell is the trade winds. These winds blow nearly continuously at the earth's surface from east to west just north and south of the equator in the Pacific, Atlantic, and southern Indian oceans. How does that work? Hadley Cells tend to make air flow toward the equator, but the Coriolis Effect (sidebar 1) bends the airflow toward the west (right in the Northern and left in the Southern Hemisphere). The result is continual airflow (trade winds) that averages a bit over 10 mph, although they can be stronger. The third result is that the air movement higher in the atmosphere (roughly 8 miles up) also experiences the Coriolis Effect, which makes it bend right (east in the Northern Hemisphere) producing the Subtropical Jet Stream. This jet stream can exceed 100 mph! Why so much faster than the trade winds? One reason is that the trade winds encounter more friction in passing over the water; another is that the air is less dense at higher altitudes. Another part of the story is that a stronger pressure gradient aloft drives stronger winds. In this case the meridional (or north-south) thermal gradient (associated with the pressure gradient) is stronger at higher altitudes than at the surface.

In contrast to the Hadley Cell, the Polar Cell is driven by cold, dense (heavy) air that descends over the poles and pushes air toward the equator. As the air moves toward the equator it warms, becoming less dense and moister until at about 60°N (or S) it rises and again moves toward the pole. As it cools, the water condenses into precipitation: rain in the summer and snow in the winter. The effect is certainly noticeable in Juneau (58°N), where I live, where average rainfall is about 62 inches as compared to 50 for New York City, 37 for Seattle and Chicago, and 13 for Los Angeles.

Between the Hadley and Polar Cells, air moves toward the poles along the earth's surface, and the cycle of this Ferrel Cell is completed by movement at higher altitude. Again the Coriolis Effect bends the...
winds toward the east (right in the Northern and left in the Southern Hemisphere). A polar jet moves air to the east near the high atmosphere intersections (4 to 8 miles) of the Ferrel and Polar Cells. You can look at the Ferrel cell as a gear that connects the Hadley and Polar Cells.

and at those times the house is assaulted by winds from the north that reach 60 mph and hit the house hard enough to cause it to move and make waves in our little cove and even our toilet bowl! Those winds are the result of the clockwise movement created by the high pressure system, in which the air moves down and out from its center and is influenced by the Coriolis Effect to produce a pattern just the reverse of that shown in Figure 13.

Effects of the Moon
The moon is primarily responsible for producing tides on earth (sidebar 3). Tides are responsible for large movements of water, particularly in coastal areas. Advance and retreat of the tides continually renew important components of the nearshore environments, by removing and replacing water that replenishes oxygen and food and by maintaining temperatures. Strong tidal currents occur in locations that have landforms that channel water. The stage of a tide (e.g., high or low) can determine whether or not a boat can safely navigate, if it is possible to dig clams or rake crabs, or when small salmon should descend from its stream to the estuary. Most of the small salmon with which I work emigrate to the ocean in mid April. Why? At that time there is a new moon and predators have a harder time detecting them in the dark. In addition, a high tide (sidebar 3) shortens the distance they must travel to salt water.

Fitting the parts of the puzzle that influence GOA oceanography
So I hear you yelling: “Get on with the Gulf of Alaska Oceanography!” Well, it’s not quite that simple. There are a number of factors (pieces to the puzzle) away from the GOA that influence GOA currents and water movements. So let’s pick up where we ended when we looked at Hadley Cells (sidebar 2). Remember that one result was that currents near the equator flow from east to west. The currents we will describe are persistent over decades, although their magnitude may change because of weather patterns. Since the GOA is in the Northern Hemisphere, we’ll follow the current north of the equator, the North Equatorial Current (Figure 14). The North Equatorial Current approaches land near the Philippine coast and splits into two branches, one that flows north and one that flows south. The northerly flowing current, the Kuroshio Current, carries warm water past Taiwan and along the southern Japanese coast. Here, the warm Kuroshio water collides with the cold Oyashio Current and bends east across the northwestern Pacific Ocean (where it is called the Kuroshio Extension). The Oyashio Current originates from the southern tail of the East Kamchatka Current after part of it has taken an excursion into the southern Sea of Okhotsk. The Oyashio Current splits near the south end of Japan’s Hokkaido Island. Part of it moves southwest along the east coast of Japan where it collides with the Kuroshio Current and swings east to form the Subarctic Current. At about the International Date Line (180°), the Kuroshio Extension and Subarctic Current meet and flow east parallel to each other to form the broad (40-50°N latitude) North Pacific Current, which flows to the northeast of Hawaii and continues east to the Pacific coast of North America. Note that the winds from the atmospheric cells and both landforms and ocean bottom topography (bathymetry) determine these paths. If you need proof that there is a current that transits the North Pacific, take notice of the debris from the 11 March 2011 tsunami in Japan that has washed ashore on the coasts of Hawaii (a boat), Oregon and Washington (floating docks), and British Columbia (a motorcycle) (http://www.huffingtonpost.com/2013/03/09/japan-tsunami-debris-earthquake-photos_n_2843978.html#slide=1163971).

The water from the Kuroshio Current is warm and salty, whereas the Oyashio water is cold and has low salinity because it is carrying runoff from the land, the heavy precipitation associated with the
What Does Genetics Have to Do with It?

THE TIDES

My grandfather did highway maintenance in the state of Washington. One day when he was working near some tidal marshes, a flatlander (an Ohioan I think) stopped, got out of his car, and asked where the water he had seen in the slough the previous day had gone. My grandfather politely told him that he had seen the water at high tide and it was now low tide. The man did not believe him; he’d never heard of tides. I wonder what he thought had happened to the water. Tides are familiar to coastal residents, and most of us know they are in some way correlated with the phases of the moon because tide tables include that information (look at the at bottom of Figure 7). Tide tables also often include a prediction of the probable fishing success (which may be related to the ease of keeping the bait on the bottom where the fish live) by showing fish drawings (halibut in our area) of different sizes.

Tides, which are the rise and fall of sea level, can be predicted at any location on earth. A number of things influence the tides. The most important is the moon but other things such as the earth's rotation, the sun, the shape of the ocean bottom (referred to as bathymetry), and the isolation of parts of the oceans by land masses (continents and islands) all play roles. Let's do something scientific: construct a conceptual model to explain tides, starting with the most significant factor and adding others until the model is predictive of observed processes.

We will begin the process with the moon, the most important factor (the moon accounts for about two-thirds of the total effect of the sun and moon). The earth and moon attract each other because of gravitational force (recall Isaac Newton and his apple). Although we think of the moon orbiting the earth, they actually orbit each other. This means that the earth is not a fixed point around which the moon rotates. Both rotate around a point called the center of rotation, but they are not equal distances from it. That means that not only does the earth spin on its own axis, it also moves as a body around the moon. Here is an example that will demonstrate why the center of rotation is closer to the earth than the moon. Let's consider the hammer throw (one of four throwing events in track and field). The hammer is a 16 pound weight...
attached to an approximately 4 foot chain. The athlete spins around holding onto the handle and then releases it tossing it as far as possible. In Figure 8 an athlete in mid throw is leaning away from the weight. The actual center of rotation is between the athlete and hammer, closer to the athlete because of his greater mass. Similarly, the center of the earth/moon pair is much closer to the earth. Why is this important? If the earth and moon did not orbit each other, they would attract each other and the moon would crash into the earth, like Newton’s apple and the ground, causing mass extinctions of life on earth.

Remember, our goal is to understand what generates tides. Unlike the hammer, the earth is not completely solid; it is surrounded by a gaseous atmosphere and has an abundance of liquid on its surface. Because water and air are fluids, they can be distorted by the moon’s gravitational attraction. Think of swinging a water balloon around on a string. The force of the string will draw one end of the balloon toward the string, but the other side will pull in the opposite direction (centrifugal force from the inertial tendency of the water to keep going outward, like a child flinging off of a merry-go-round). The balloon will distort into an elongated shape (Figure 9). Similarly, the earth’s waters bulge toward and away from the moon depending on where the water is in relation to the moon (Figure 10). The bulge away from the moon is not quite as large as the one toward the moon. The two bulges mean that in many places on earth there are two tidal cycles per day.

If the earth were smooth and completely covered in water, we would be finished considering the effect of the moon on tides. But there are a few other things to consider. One is that the moon orbits the earth in the same direction that the earth rotates, just not as fast. The earth completes a rotation in 24 hours. However, during that time the moon has also moved so it takes about 24 hours plus another 50.5 minutes until the moon is over the same place on earth that it was the previous day (Figure 11). This means that the tides do not have a precise 12 hour cycle (Figure 1). Because the water and the solid earth on which it sits are not frictionless and water has inertia (sidebar 1), the water tends to rotate with the earth. As a result, the bulge of water is not directly below the moon; its position lags a bit behind the moon’s position. Note on the tide table that the very high and low tides are on June 24 and 25, not on June 23, the day of the full moon (Figure 7). Also note that there was another high tide series earlier in the month near the timing of the new moon. These two moon phases
What Does Genetics Have to Do with It?

Sidebar 3 continued

The earth is tilted relative to the sun and the moon is tilted relative to the earth’s orbit around the sun. As the earth rotates, the same location about 12 hours apart (A and A’) will experience different tidal heights because they are under different parts of the tilted tidal bulge.

The next factor to incorporate in our model is that the earth is tilted relative to its orbit around the sun (sidebar 1) and the orbit of the moon is tilted about 5° relative to the path the earth tracks around the sun. Consequently, the moon does not orbit directly over the equator, but on a path that may swing as far as 18.5° north and south of the equator. The track of the water bulge encircles the earth like a sash worn by Miss America or royalty. As the earth rotates, some locations will be under a deeper portion of the bulge on one side during the daily rotation (A in Figure 12) than on the other (A’). That explains the difference in the magnitudes of the high (and low) tides that occur twice daily (Figure 7).

The final important factor that affects tides is the shape of the land and sea bottom. Clearly, the tidal bulge cannot pass through land barriers. Several major bodies of water are surrounded by land or by a combination of land and shallows that alter the effects of the sun and moon—some examples are the Bering, Mediterranean, and Caribbean Seas. Features can result in virtually no tidal range in some locations, like the Gulf of Mexico, and extreme tides in other areas, like Canada’s Bay of Fundy (53 feet—don’t get caught in an incoming tide!). The very high and low tides—the times when the bulges of water are largest—occur when the sun and moon are aligned and “pulling together.” Also, the geographic position and land and seafloor shapes can locally accentuate or dampen the effects of the sun and moon. Complexities such as the shape and position of the land and sea bottom are difficult to model, but scientists continue to improve models for local areas.

We have seen that the moon and the landmasses on the earth have major effects on the water tides. It also turns out that there is a measurable “tide” affecting the landmasses and atmosphere! The gravitational attraction that produces the tides is not just between the earth and moon/sun, but is between each and every molecule on the earth and moon/sun, including all the molecules that form the solid continents. Although the solid parts are pretty well (but not completely) fixed in position, the earth actually gets stretched a bit (measured in centimeters) by the moon’s gravitational attraction. In contrast, the atmosphere can have tides that are measured in kilometers.
Aleutian Low over the North Pacific Ocean, and sea ice meltwaters exported from the Sea of Okhotsk. As they flow to the east in the North Pacific Current, they remain separate. Their interface creates a sharp north-south change in both temperature and salinity called the Subarctic Front and is carried across the Pacific. The Subarctic Front forms a relatively stable boundary that marks the southern edge of the Gulf of Alaska. The North Pacific Current splits into two branches near the North American coast. One branch (the California Current) flows south and the other (the Alaska Current) flows north. Both of these are wind driven. The California Current is influenced by winds resulting from the Hadley Cell (sidebar 2), and the Alaska Current is driven (particularly in the winter) by the counterclockwise winds of the Aleutian Low.

The Gulf of Alaska

The Gulf of Alaska is the large portion of the northern Pacific Ocean that is bounded by the central Alaska coast and Aleutian Islands in the north and by Southeast Alaska and British Columbia in the east. Mountain chains that rise more than 3,000 feet extend from British Columbia to the Aleutian Islands. We will see that these mountains strongly influence the oceanography of the GOA. To the south the GOA is bounded by the Subarctic Front, while to the west the GOA boundary is less well defined.

The circulation of water (currents) around the GOA is primarily wind driven. The atmospheric cells (sidebar 2) and weather patterns (e.g., the Aleutian Low) drive the currents in a counterclockwise direction. It’s time for a look! The North Pacific Current bends north to produce the Alaska Current (Figure 16). The Alaska Current (AC) lies beyond the continental shelf break and flows at a leisurely rate along the eastern GOA (about 4 to 13 km per day) and it may be 250 km wide or more. In the northern GOA west of about 150°W, the AC becomes the Alaska Stream, which is much narrower (about 75 km) and faster (about 43-86 km per day). The continental shelf, which is the coastal boundary for the AC, can be narrower than 20 km in the southeastern GOA but it extends more than 200 km from shore along the northern GOA. The AC often

Figure 13. Left: Anticyclonic (counterclockwise) rotation of Northern Hemisphere low pressure systems. Air distant from the low pressure air (blue arrows) rushes toward it and into it but is deflected to its right (red arrows) as a result of the Coriolis Effect (sidebar 1), causing the system to rotate counterclockwise (black circle). From: https://en.wikipedia.org/wiki/File:Coriolis_effect10.svg. Right: NASA photograph of a low (anticyclonic pressure system over Greenland. From: https://en.wikipedia.org/wiki/File:Low_pressure_system_over_Iceland.jpg.
What Does Genetics Have to Do with It?


flows near the continental slope in the southeast but in the north is somewhat farther off shore but continues to hug the continental shelf until it turns into the Alaska Stream, which moves rapidly along the Aleutian Islands. Some of the Alaska Stream penetrates passes between the Aleutian Islands at several locations and feeds a counterclockwise current system in the Bering Sea (Figure 17). Following the incursion into the Bering Sea this current system coalesces into the southwest-flowing East Kamchatka Current along the western boundary of the Bering Sea, whereupon it will eventually feed the Oyashio and then the North Pacific Current to complete the circuit.

Another GOA current is the nearshore Alaska Coastal Current (ACC), which hugs the coastline over the continental shelf from British Columbia in the south to the Aleutian Islands in the west (Figure 16). The ACC is driven by wind and freshwater runoff and generally extends less than 50 km from the coastline. Its speed can approach 90 km per day in the fall and winter when freshwater discharge is large and winds are strong, but it is much weaker in spring and summer. Much of the land bordering the GOA is the recipient of the precipitation produced by the Arctic Low (sidebar 2). In addition, the surrounding mountains substantially increase the amount of rain and snow. How? The winds that pick up moisture from the ocean and blow from the southeast encounter the mountains and are forced to rise. As the
air ascends, though, it cools and releases most of its moisture before passing over the mountains (Figure 17). The air that ultimately passes over has lost its moisture to rain and snow and the other side of the mountains is often much drier. Also, because the air is not tempered by the high specific heat of the ocean water, which warms the atmosphere in the winter and cools it in the summer, the interior climate is also more extreme—often referred to as a continental climate. Figure 16 includes bars along the coast that show average annual precipitation. Note that 2 meters per year is nearly 80 inches and many locations receive more than that—Yakutat receives 169 inches of precipitation per year!

In contrast to the large, relatively stable Alaska Current and Alaska Coastal Current, large clockwise eddies form along the continental slope between British Columbia and Yakutat. These eddies typically originate annually in winter. Eddies that form off British Columbia usually move across the interior GOA basin, while those that form between Sitka and Yakutat generally hug the continental slope and move around the GOA. The eddies are large (about 150 km in diameter) and deep (they extend vertically at least 2,000 m). They can persist for from 2 to 5 years, while moving slowly (about 2-3 km per day). During spring and summer at some times and locations, the eddies can impinge onto the continental slope and shelf break and alter outer shelf currents. Although the eddies themselves are moving slowly, the currents within the eddies may be quite vigorous and have swirling velocities of 20 to 40 km per day. In Southeast Alaska the edge of the eddies may even reverse direction of the ACC. These eddies are ecologically important: they bring into contact water from rains and melting snow and ice, which picks up the element iron as it flows over the land, with nitrogen-rich water
from the GOA basin to create a rich environment for phytoplankton, zooplankton, and ultimately fish. The eddies can be detected by remote sensing of the chlorophyll that is produced by phytoplankton in those nutrient-rich water mixtures (Figure 19).

There are also much smaller (about 20-30 km wide) eddies that form on the continental shelf. They are generated through interaction of the coastal current with the rough bottom topography or upon flowing past something that sticks out and alters the flow, such as coastal headlands like Cape Edgecumbe near Sitka and Kayak Island near Yakutat. Underwater canyons (e.g., Amatouli Trough and Hinchinbrook Canyon) may alter water and large intrusions from freshwater runoff (e.g., through Cross Sound and from the Copper River) may be involved. Finally, wind reversals, such as may occur as a result of intrusions of arctic cold fronts in the winter may alter water flows.

In addition to the many roles the Coriolis Effect plays in oceanographic features that we have reviewed so far, there is one more behavior of water that is critical for coastal organisms: upwelling and downwelling. Along the GOA coast, the wind blows counterclockwise. These winds push surface water to the right of the winds and toward shore. This phenomenon, known as the Ekman effect, is due to the Coriolis Effect in conjunction with frictional forces in the water. As the surface waters move onshore they raise the sea level and force nearshore water downward (downwelling) at the coast.

In areas where the wind blows to the south, such as where the Oyashio and California Currents occur, upwelling results because the wind blows surface waters offshore, lowering coastal sea levels and bringing deeper water to the surface near the coast. Upwelling carries nutrients from the bottom and makes them available to organisms closer to the surface. As a result, areas of upwelling are often very rich in the diversity of sea life, which explains why those coastal areas are often biologically richer than open ocean areas.

Summary
The description of Gulf of Alaska oceanography that has been presented suggests a relatively static situation. Do not interpret it as such! What we really have done is provide a beginning. From that we have to include variations that occur in time frames from minutes to decades (or more if you are interested in evolutionary changes). We all know that even though science underlies weather forecasts, the weathermen in some regions of the country (Juneau for instance) are not always accurate. So, look at this description as a theme (as in music) and appreciate the variations on the theme that time presents you. And if you are in Yakutat, do not forget your boots and raincoat!
It is often difficult to know just what features of the environment produce population structure in marine species (recall Chapter 2. How Genes Vary in Fish Populations, and Chapter 5. Fish Population du Jour). We can measure currents and observe geologic features that we think might influence their movements, but unless we mark and track individuals, we cannot be sure if they travel extensively or are homebodies. In this chapter, we consider the implications to conservation and management of the extent to which individuals move around. We also look at the genetics of a northern Pacific Ocean rockfish species, Pacific ocean perch, to see what we can learn about their movements in relationship to the oceanographic features they encounter in the Gulf of Alaska.

Recall that either behavioral or geographic separation of groups of individuals of a species can produce populations that differ genetically (Chapter 2); that is, there is population (genetic) structure. If we detect genetic structure in a species, we can use it in reverse to learn about the biology of a species.

One of the most important aspects of fisheries conservation and management is the ability of a species to produce more of that species. The number of fish that can be harvested, without depleting the stock, ultimately depends on the ability of the species to replace harvested fish. Managers use stock assessment to quantify the numbers of fish available. Stock assessment estimates the numbers (or biomass in tons) of a species within a geographical region that may be fished. Stock assessment surveys generally sample specific, predetermined locations to estimate the abundance of fish. Often they conduct surveys.
periodically (bi- or triennially) to estimate trends in abundances. Population dynamics describes the ability of and rate at which a population produces more individuals in the face of natural (e.g., old age and predators) and fishing mortality. Population dynamics ordinarily involves the application of sophisticated mathematical models that incorporate stock assessment, harvest, life history, and even oceanographic and environmental information. Federal and state agencies devote considerable effort to such activities.

Population genetics studies have only recently begun to be used to improve conservation and management. It turns out that the genetic structure of a species can provide information about both the biology and sustainability of harvest practices. One of the things that we can learn from patterns of genetic differences observed among samples of a species that are collected at multiple locations is how populations are distributed over a geographic range. If the species is broken up into discrete populations, genetic information can provide information about how many populations there are or how far apart they might be. Even if a species is distributed fairly evenly over a large area (like the Bering Sea shelf), genetic differences may exist among collections sampled at different locations. How can this be? If larvae, juveniles, and adults do not travel (disperse) very far between birth and reproduction, mating will occur only between individuals in a local area. This will produce more of a patchwork of groups than a single large group. Genes from a particular group will only travel as far as the fish that carry them. Even if there are no obvious geographic, oceanographic, or behavioral boundaries separating the groups, the limited distances traveled during a lifetime (dispersal distances) will prevent genes from spreading widely. If the dispersal distance is large, there will be little change in genetic composition (genetic divergence) over long distances. In contrast, if the dispersal distance is small, families (and their genes) will be restricted to fairly small areas. The concept is easy to understand when one considers salmon populations because salmon return (relatively) faithfully to their natal streams to spawn. Isolated populations can diverge genetically as a result of random drift and local selective pressures (see Chapter 3. History of a Salmon Population). Species that are distributed fairly evenly (often referred to as continuously) over a large region can also diverge from area to area if they do not move very far between birth and reproduction and, therefore, do not exchange genes very often. It might be easier to think of species that do not pick up and move, like plants. In some mountain meadows, you might see grass over most of the meadow, but wild iris might be in patches. Let’s look at the iris: at one edge of the meadow there may be normal purple iris, but in another place they may be white. Clearly, each patch shares a color gene. Pollen can be blown or carried all over the meadow (or to other meadows), but most of the successful pollination occurs close to the donor. Even though fish can swim long distances, they often do not, because their behavior keeps them close to home, where historically there has been food and habitat that can sustain their kind. The management implication of such structure is that many sites of production of the species may be spread over the species range, and uneven patterns of harvest can reduce production as a result of a patchwork of overharvest and underharvest, neither of which is optimal for sustained production.

For fisheries directed at species like salmon, the importance of taking spatial structure into account in management plans is obvious. Salmon ordinarily return to their natal grounds to spawn. Their homing precision, although not always perfect, is good enough so that if a spawning run were severely depleted or wiped out, say by a net or trap set across the mouth of the stream, it is likely that it would take some time to reestablish and rebuild from strays from other streams. In addition, some drainages have more than one run (spawning aggregation = population) of a species. Different runs often spawn in areas within a drainage that differ environmentally from each other. For example, different tributaries may have different temperature regimes or water flows. In other instances, spawning habitat may be limited and temporal spacing of runs may increase the seeding density. In both instances, each run encounters somewhat different environmental conditions, which impart different (natural) selection pressures on the populations. It is likely that the runs will respond to local selection differences and become locally
adapted; that is, they will diverge genetically for traits that are important to productivity in each particular environment (see Chapter 3. History of a Salmon Population). Even for these geographically proximal streams, if one of the runs were removed, it could take some time to restore. So for salmon, it is clear that temporal and spatial population structure must be considered in the development of management plans.

Consideration of spatial and temporal structure is important for marine species. Unfortunately, the possibility of spatial structure in marine species is often ignored or dismissed as unimportant or nonexistent—what cannot be easily seen is often ignored. For many fish species, the possibility of spatial structure is often disregarded because the fish live to be very old and there are often no obvious (to us) barriers to the movement of adults. In addition, many species produce eggs, larvae, and juveniles that have the opportunity to drift with prevailing currents. But, just because fish can travel doesn’t mean that they do. Because the potential for movement (dispersal) throughout fishes’ lives is high, many biologists and oceanographers treat the assumption of wide dispersal as fact in developing management models. Clearly, some species do travel long distances and can reach nearly any part of the natural habitat of the species during their lives. However, we are learning that many species, in spite of the potential for them to disperse, either do not move far from the places they were born (released into the environment by their mothers) or else they return to those sites to spawn. Below, we examine how population abundances may be eroded if spatial structure is ignored. We then briefly examine the results of our genetic study of POP (Pacific ocean perch, *Sebastes alutus*), which illustrate that POP have strong spatial structure even though they have the potential to travel long distances.

To examine the potential effects of ignoring spatial structure in managing a species, we will consider two strongly contrasting models. To keep it simple, we ignore any potential local adaptation (see Chapter 1. Even Fish Obey Mendel’s Laws) that may result from spatial structure (but if it exists, it would increase the extent of genetic divergence over space). Both models describe species that are broadly distributed (Figure 1). These models consider relative sizes of geographic areas (spatial scales) that are specifically defined in fisheries management regulations (for example an embayment or portion of the Gulf of Alaska) and over which a species actually travels between birth and reproduction (dispersal distance). The first model describes a species for which the area defined for management (management scale) is less than the average distance that individuals travel between from birth to reproduction (dispersal scale) (Figure 2 left). The second model describes a species for which the management scale exceeds the dispersal scale (Figure 2 right).

In the first model fish are removed by a fishery, but the larger dispersal distances would allow the spatial distribution to even out rapidly (Figure 3). In the second model, however, the small dispersal distances would leave the distribution sparse in some areas and unchanged in other areas (Figure 4).

A simple approach to population dynamics is to examine the relationship between the number of spawners (often females because the number of eggs is the total reproductive potential) to the number of recruits (the number of fish that enter the fishery) that they produce. Often the relationship is graphed over time and plotted as stock-recruit curves (sidebar 1). If we consider the consequences of our two models within the framework of that standard, we see that model 2 has the potential to
erode productivity. Figure 8 presents the stock-recruit curve and exploitation target of a well-managed fishery. The curve plots the number of recruits (on the \( y \) axis) that are produced from the abundance of parents (the \( x \) axis). The diagonal line plots replacement—the number of recruits equals the number of parents. Most populations can produce more than one recruit per spawner at lower densities. At higher densities—to the right of where the curve intersects the diagonal line—some form of resource limitation such as food or spawning habitat reduces production to less than replacement. If there is too little food many individuals will starve; if there is too little spawning habitat, there will not be enough space for all of the females to spawn. The vertical line shows the number of spawners that produces the largest
One of the things that fisheries scientists do is project how many fish will be available for harvest. They have the responsibility for balancing harvest demands with conservation goals. This is a challenging job because harvesters usually want to catch as many fish as possible, although the market, fishing effort and costs, and other factors may influence them. On the other hand, sufficient breeders must survive the harvest to perpetuate the stock. This is the high-wire act of managers. How do they do it?

Let’s think about how this might be done. One of the fundamental ideas is that big fish make baby fish. So, it should be clear that if there are no adults, there will be no babies. This idea might lead one to think that, if we just let a population go for a while, we could eventually obtain unimaginably large numbers of babies. But that conclusion would be wrong. Why? Think about it. If we let the population go for a very long time, it would run out of food, places to live, places to spawn (especially salmon, which dig redds [= spawning nests] in streams), and so forth. These things that would ultimately limit the size of a population are referred to as resources. And as a population size increases, eventually it reaches a point at which the resources are completely used up. Resource limitation is one of the reasons that aquaculture is applied so widely. An aquaculture facility can artificially provide resources like food, spawning habitat (in the form of incubators), or space to grow (in the form of raceways) (see Chapter 6. Genetics and Hatcheries). Because at some point resources will limit the population size, the relationship between the numbers of parents and offspring is not the simple linear relationship that we might have naively assumed before we did some thinking.

Let’s look at population size in a different way. Maybe we can find a better way to relate the numbers of parents and offspring. We’ll start small with a bacterium. Again, the process we’ll follow is an oversimplification, but it will get us started. We will (conceptually) inoculate a single bacterium into a flask that contains growth medium (i.e., food). The bacterium will grow in the medium and divide into two bacteria. Those bacteria will in turn grow and divide into four bacteria, which will make eight bacteria, and so on. It is amazing how rapidly the population size grows (Figure 5A). However, as the number of bacteria increases, they consume the available food. Over time, the bacteria population grows slower and slower until all of the food is gone. At that time, the bacteria will stop growing and dividing (Figure 5B). Similar limitation on food can also regulate populations of other organisms. In addition for some organisms, the amount of space available (say for spawning) may limit the size of the populations. The maximum population size that can be maintained by a stream or reef or other discrete habitat is referred to as the carrying capacity. The curve that describes that restrained growth is “S-shaped” or sigmoid. Bacteria do not necessarily die immediately, but the population size reaches a maximum size. In contrast, most fish eventually die and they can live only so long without food (although some can survive for very long times).

We also can use the graph in Figure 5B to describe growth of fish populations; we just need to change the concept a bit. A variety
of resources may limit the size of a fish population. For fish, some resources like spawning habitat are more or less fixed. If a key resource (like spawning habitat for salmon) is exhausted by demand, an excess of another key resource (like ocean rearing capacity) cannot compensate. The limiting resource will determine the maximum population size. Resources like food can be limiting, but they are more complex because the food resource itself may be regulated. For example, zooplankton is a critical food source for young salmon and other species; their abundance is regulated by phytoplankton abundance, which is regulated in turn by sunlight and other resources. Both zooplankton and phytoplankton abundances are renewable, but their population abundances can also be described by “S-shaped” curves. This means that if food is the limiting resource, a fish population will be regulated by a series of production curves for the species in the food chain. These kinds of interactions are what ecology is about. We won’t consider these relationships any further.

Now let’s return to our consideration of the relationship between the abundances of parents and offspring. It should be obvious that a female must contribute two offspring to the next generation of spawners to replace her and her mate. What allows fisheries harvests to be possible is that at moderate abundances (relative to the carrying capacity) each adult female may contribute multiple offspring to the next generation. Note that we generally focus on the numbers of females. Except when they are at really low abundances, the number of males have little influence on the productivity of populations. That is why we have male-only crab and deer harvests. As a population size approaches its carrying capacity, the average numbers of offspring that are contributed by each spawning female decreases. One of the most common ways to visualize this relationship is with a Ricker plot, named after the eminent fisheries scientist William Ricker.

The Ricker plot shown in Figure 6 shows a hypothetical relationship between the number of spawners in two consecutive generations, parents and their offspring. This diagram has three lines. The diagonal blue line (a) is a reference line that shows how many offspring must be produced to exactly replace the parental population. Points along this line represent an average contribution of one offspring by each individual (or two per female on average). The domed green line (b) traces the actual relationship between numbers of parents and offspring. Points along the green line (at the left) that are above the blue replacement line indicate that the population is growing. Points along the green line that are below the replacement line (at the right) indicate a declining population.

The vertical red line is what makes harvesting possible. The red line (c) represents surplus production. That means that the number of fish represented by the red line can be removed (harvested) from the population prior to spawning and the number of spawners will be the same as it was the previous generation. This means also that a harvest of that size can be made every generation without eroding the population. One could search all of the vertical lines that extend upward from the blue line to the green line to find the single largest line. Maintenance of that number of breeders would generate the maximum yield for this relationship.

Sometimes different stages in the life history are used, but the stages represent consecutive generations of the individuals considered. For example, in some instances the number of fish recruited to the fishery or the number of returns to a stream in the next generation per adult spawner is plotted. Of course, the plots must be adjusted if a species matures at more than one age (like Chinook salmon) or spawns multiple times.
One word of caution is that the green line is ordinarily estimated from multiple years of data for a population. The line itself results from fitting the population size generation $t$ on the $x$-axis to the size at generation $t+1$ on the $y$-axis. There is usually substantial variability in the response of the number of offspring to the number of parents because of climate differences between generations and other environmental influences. In addition, there are usually not very many points to use in plotting the curve. As a result, the information should be used very carefully. Another caution is that the actual parameters that underlie the curve (like carrying capacity) may vary over time. For example, global climate change may gradually increase or reduce the carrying capacity. In fact, interannual climate variability is one of the reasons the data used to plot Ricker curves is often so chaotic looking (Figure 7).

![Figure 7](image-url)

Figure 7. An example of a Ricker curve for the relationship between the number of spawning pink salmon females and the number of fry that emigrate from Auke Creek the following spring. Note that young pink salmon leave the creek soon after they emerge from the gravel and only a small portion actually return to the creek to spawn. Consequently, the fry survival reflects the entire freshwater experience, from egg deposition through incubation to emigration. Note that even though the plot looks chaotic, there is a very good fit to the curve (the chances of observing a worse fit to the data is only about 2 in a million). The arrow indicates the optimum number of females.

The number of recruits per spawner. The portion of that line that is above the diagonal line represents the surplus production, that is, the maximum proportion that can be caught without reducing the productivity of the system. A goal of management is to determine these values and manage harvests and stock sizes accordingly.

In model 2, the reduced dispersal distance means that there will be areas that are not at optimal densities (under- or overutilized habitat). As a result, the densities of parents will either be lower or higher than the density that optimizes production in Figure 8, and the overall production of the management area will be reduced (Figure 9). If the dispersal distance is quite small, persistent fishing can severely deplete the resource in some areas. For species like POP that are very long-lived, recovery could be very slow.

The very simple spatial models we presented assume that POP occur all along the edge of the continental shelf of the Gulf of Alaska (distributed continuously). In reality, POP are found at many places along the shelf edge, but not every place. They tend to cluster near marine valleys that lead from the shelf edge to deeper water, where oceanographic processes concentrate food. Stock assessment is challenged by the patchiness of their distribution, and special sampling methods have been and continue to be developed that take into account the patchy nature.

Now let’s take a look at what we’ve learned recently about POP (remember that these are Pacific ocean perch). Our work has been supported by NOAA Fisheries, North Pacific Research Board, and Alaska Sea Grant. Our objective has been to describe the genetic structure of Alaska POP and to use that information to learn about their life history. Even though POP have been harvested for more than a century, there are many things we still do not know about them.
So what do we know about them? Pacific ocean perch are distributed along the Pacific Rim from California to Japan. They have been fished heavily throughout their North American range since the 1940s and are a valuable component of the Alaska groundfish fishery and probably the marine food web (see Chapter 10. The Wonderful and Diverse World of Rockfishes). POP are most abundant in Alaska waters, inhabit both the Gulf of Alaska and the Bering Sea, and are the dominant species among the rockfishes that are found along the continental slope. POP are slow-growing and long-lived—some live more than 100 years—and they have a relatively late age at 50% maturity—8 to 10 years or so (50% of the fish are sexually mature at age 8 to 10). They recruit into the fishery (become catchable) between ages 7 and 15, which introduces a lag in observing the effect of overfishing on the reproductive potential of the parents of these recruits. Fishing pressure also can reduce the reproductive potential of stocks by reducing the age at maturity. In the northern Washington–Vancouver Island area, ages of POP stocks at 50% maturity were reduced from 10.1 to 8.1 years between 1972 and 1992.

Many aspects of their life history are still unknown, and what is known is based mostly on observations of populations in the Queen Charlotte Sound, which is a bit more temperate than the Gulf of Alaska and Bering Sea. Like other rockfishes, POP are live bearers (see Chapter 9. The Challenges of Taxonomy as Seen through the Eyes of Rockfish Biologists), and females release planktonic larvae (they drift passively in the ocean) that are ready to feed. The locations at which insemination (transfer of spermatozoa from males to females), fertilization (fertilization of the eggs), and parturition (release of larvae) occur are unknown, but may be geographically separated. Insemination takes place sometime in the fall, but spermatozoa are retained until fertilization occurs in deep water, probably during the winter. The release of the larvae in April and May probably coincides with plankton blooms in the Gulf of Alaska.

Little is known about dispersal and movement after the larvae are extruded by the females. Juvenile POP are released to a planktonic life history phase, and the larvae probably settle to the bottom (become demersal) within the first year of their lives. Larvae may remain in the open water (pelagic) for only several weeks or may remain off the bottom for a few months. But during that time, their movements may be influenced by prevailing currents, which could play an important role in the variability of year-to-
year recruitment. Although little is known about the larval or early juvenile stages, 18-month-old juveniles appear on the continental shelf off British Columbia. Also, 1-6 year old fish have been seen at the bottom of coastal fjords in Southeast Alaska, where they feed mostly on small planktonic or pelagic crustaceans. In Southeast Alaska, the segregation of young POP by age generally corresponds to depth because older fish are often found in deeper water. By age seven, the Southeast Alaska POP usually recruit (become a member of) to adult stocks offshore, even though they may not yet be sexually mature.

The planktivorous POP are generally semi-demersal, but can be pelagic at times. Adult POP are usually found in water between 90 and 825 meters (m) deep; however, both males and females are most often at 200 to 275 m in the summer and 300 to 450 m in the winter. Females may move deeper (500 to 700 m) after insemination, a depth at which we think they remain until their larvae are released.

It sounds as if we know a lot. So, what don’t we know? We do not know precisely where the fish are at insemination, fertilization, or release of larvae. We also do not know how far adults wander or if they return (home) to specific locations for reproduction, feeding, overwintering, and other activities. We do not know to what extent larvae are influenced or transported by currents such as the Alaska Coastal or Alaska Currents (see Chapter 12. It’s a Dangerous World Out There! Oceanography of the Gulf of Alaska). We do not know how far larvae and juveniles drift. In fact, extreme movements may carry a lot of them out to sea where many perish. Consequently, successful capture of small fish documents their existence at a location, but it does not mean that fish from that location could eventually recruit to the adult population.

To answer some of these questions, we conducted studies that used population genetics to examine POP population structure. From those studies, we hoped to be able to make some deductions about the extent to which POP dispersed over their lifetimes. We began the first study by obtaining tissue samples from adult fish that were collected along the continental shelf break of the Gulf of Alaska and Bering Sea during stock assessment surveys by NOAA Fisheries. The collections we examined were sampled about every 500 km along the continental shelf break from southern Southeast Alaska to the western end of the Aleutian Islands and in the Bering Sea (Figure 10). We conducted genetic analyses on the samples to obtain allele frequency data for 14 microsatellite loci (see Chapter 4. Molecular Tools for Population Genetics). The first question we asked was “Do the distributions of allele frequencies vary among collections?” The answer was a resounding yes. Every collection differed from every other collection, not just for the aggregation of results over all loci but for individual results for 13 of the 14 loci. Some caveats are necessary in the interpretation. First, although the sizes of the samples were adequate for analysis, the number of collections was limited. Many places between the sites that we sampled also have POP. Our conclusions were extrapolated to include those other locations. Second, POP can have long lives and the collections of POP that we analyzed include many different ages of fish. One might conduct a comparable study on humans by randomly sampling from multiple neighborhoods distributed throughout a large city to make conclusions about the nature of the city. Clearly, the more neighborhoods sampled, the better the conclusions. We can conclude, however, that each of the POP collections represents a distinct source of POP and that there is not much exchange between these sources, or they would have similar genetic compositions.

The next question we asked was: “Is there a relationship between the amount of genetic divergence between the collections and the geographic distance that separates them? The idea that underlies this question is that if POP do not move very far between birth and reproduction, the opportunity for their genes to spread is limited. As a result, collections that are close to each other would be more similar genetically than collections that are far from each other. The existence of this kind of pattern is referred to as isolation by distance. Again, the result was an unqualified yes. A plot of genetic divergence against geographic distance showed a strong positive relationship (Figure 11).

The conclusion from this study was that POP do not move substantial distances between birth and
reproduction. In fact, our data suggest that most POP reproduce much less than 500 km (the approximate distance between collection sites) from where their parents conceived them. It is amazing that the larvae are released where there may be a possibility that prevailing currents can sweep them relatively large distances. In addition, the adults are mobile and during their very long lives have the ability to swim long distances. Our results suggest that POP remain remarkably close to where they originated and that the geographic scale over which our samples were collected greatly exceeded the scale of population structure. Our observations are consistent with the idea that even though there may be no geographic or oceanographic boundaries separating distinct populations, limitations in the distance the fish move during their lives prevents them from mating with individuals that are not within their movement ranges. Consequently, their genetic structure resembles a patchwork quilt, the patches of which are geographic units of individuals that have some reasonable opportunity to interbreed. What is interesting about this structure in the marine habitat is that there may be a (nearly) continuous series of overlapping populations (patches). In this kind of model for population structure, there may be no actual delineation of distinct populations; however, the scale of the population structure would be critical in defining management areas as mentioned at the beginning of this chapter when we considered Figures 4, 8, and 9.

One concern that arose when we interpreted our results was that many POP live very long times and produce an enormous number (millions) of larvae during their lives. That concern arises from the fact that to maintain a stable population size, only two out of those potential millions must survive (on average) to reproduce for every female (see Chapter 10. The Wonderful and Diverse World of Rockfishes). Moreover, the females usually release their larvae into a hostile environment. There are predators out there, probably including mommy and daddy, and only if the larvae are lucky will they find food before they perish. In addition, currents may sweep them out into the open ocean and they may never return. This means that there is going to be a lot of luck involved in who survives and who doesn’t. As a result of random chance, many of the progeny of just a few females may survive in a particular year, but those of others perish. Another way to look at it is that those few females may
contribute fish that year, whereas most of the females will contribute none at all. A different set of females may contribute each year and in some years there may be very few successful contributions. The differential contribution of females that is a result of these random survivals is often referred to as a sweepstakes effect. The variable survival is also why there are strong and weak year classes (also called cohorts) in some species.

The genetic implication of a sweepstakes effect is that the genetic composition may differ among year classes within a population of a species. If the genetic compositions differ drastically, random samples from different collection sites might differ just because they have different year class compositions. How do we deal with this potential problem? First we reevaluate our results. If a sweepstakes effect were an important contributor to the strong genetic differences between collections, we might expect that there would be no geographic pattern in the differences. However, we saw a very strong pattern of divergence that we suggested was consistent with isolation by distance. The only reasonable way in which we might see such a pattern if there were a sweepstakes effect would be if we had a gradient of age classes that increased or decreased from one end to the other of the range.

The next approach to resolving the sweepstakes effect issue might be to age the fish in the samples and determine if they differ in year class composition. Is aging the fish feasible? Probably not with complete accuracy. For our application, we would need to know exactly how old each fish is. Aging fish ordinarily involves interpreting ring structure, particularly in otoliths. Otoliths are ear bones. Each year, another bone layer is laid down so that they have annual rings like trees. Ages of many species of fish can be determined by counting the number of rings in their otoliths. Precisely aging POP that may be more than 50 years old would be challenging if at all possible.

OK, how else could we address our problem? Just because we cannot age adults accurately enough to identify fish of identical ages (cohorts), does not mean that cohorts do not exist. If we could figure out a way to sample a single cohort (age class) at different locations and different cohorts at the same location in different years, we could obtain genetic data from the collections to see if there were differences. If the sweepstakes effect had an influence on the genetic compositions, different cohorts sampled from the same parents would differ.

![Figure 11. Relationship between genetic divergence and geographic distance between pairs of populations (see Figure 10). For this isolation by distance (IBD) relationship, $F_{st}/(1 - F_{st})$ is a standardized measure of genetic divergence between pairs of populations, and shelf distance was the distance between populations as measured along the edge of the continental shelf.](image)
So how can we sample distinct cohorts? Serendipity is a great word. It basically means a discovery that happens through blind luck, alignment of appropriate stars, or some other such cause. For us, the serendipitous event was the capture of large numbers of young-of-the-year (YOY) rockfish in marine surveys of young salmon. The otoliths of YOY rockfish are simple enough that they can be aged! Collections of these YOY fish were made at several Gulf of Alaska locations and in several years. Even more amazing, most of the YOY fish were POP. These collections provided us with a means to evaluate the potential influence of a sweepstakes effect on our interpretations.

We conducted a variety of population genetics analyses on microsatellite data from the same loci we studied in the adult POP. The short story is that collections sampled near each other in a particular year were similar genetically. Collections sampled at the same location but in different years were also similar and not drastically different from nearby adult collections. However, collections of YOY sampled at different geographic locations across the Gulf of Alaska differed genetically, just like the collections of adults did.

The genetic structures of YOY and adult POP collections have similar isolation by distance (IBD) patterns (as in Figure 11). Apparently POP individuals do not move very far during their lifetimes; the range is much smaller than the species range. Moreover, YOY fish collected off Southeast Alaska and in the northern Gulf of Alaska from Cordova to Kodiak are probably derived from adults in close geographic proximity. Collections taken between those areas, however, appeared most similar to the western adults. Most of the collections of YOY made in the intervening area were off the continental shelf in water that was deeper than 1000 meters, an area between the Alaska Coastal Current (ACC) and the Alaska Current (AC) (see Chapter 12. It’s a Dangerous World Out There! Oceanography of the Gulf of Alaska). Both the ACC and AC flow counterclockwise around the Gulf of Alaska. This area between them is often influenced by large clockwise eddies, but the influence of those eddies varies from year to year.

Several observations of the YOY collections indicate that oceanographic features influence some YOY movements. Several transects (a set of collections sampled at nearby sites along a line in the same year) differed in genetic structure from one end to the other. Of particular note is a transect made off of Cape Yakataga, near Yakutat, which extended from the continental shelf past the shelf break. The collections over the shelf differed from collections past the shelf break. Such differences might occur if the YOY on the shelf encountered different current patterns (the ACC) from those past the shelf break, which may have been exposed to the clockwise eddies that occur between the ACC and the AC (see Chapter 12). Another observation was that the genetic compositions of several collections from transects made at the same location differed between years. Two of the sites were off the shelf in the area between the ACC and AC that may be influenced by the clockwise eddies, but the characteristics of which differ between years.

What picture emerges from these analyses? The Alaska Coastal Current is a dominant feature of the Gulf of Alaska oceanography. Parturition (release of live larvae) of POP occurs along the continental slope sometime in the spring, probably at depths between 500 and 700 m. In some locations, such as Southeast Alaska or near Cape Yakataga where the shelf is narrow, release may not be far from the ACC. Elsewhere, they are probably released away from the ACC in the area between the ACC and AC. During the time between parturition and when our collections were taken in late July or early August, the ACC is not strong because total freshwater runoff is lower and winds are not as strong as in the fall and winter months. During spring and summer at some times and locations, the ACC may even reverse direction and be underlain with protrusions of more saline water from the Gulf of Alaska basin. The consequence to larvae and young-of-the-year POP is that those that are near the surface seaward of the ACC will not be entrained in a current and even those that enter the current may not be transported rapidly or far. Consequently, the larvae may not be transported far from their release sites. In addition, several features may disrupt the current when it is
weak. For example, coastal headlands near Yakutat steer the current off shore, underwater canyons in the northern Gulf of Alaska may alter flows, and areas that channel large freshwater runoff may also influence oceanographic conditions. It is possible that these physical features produce incompletely isolated areas along the coast for POP and that the adult collections represent different areas. Clearly, additional work is needed to improve our knowledge of both the marine environment along the Southeast Alaska coast and of the biota that live there.

POP have strong population structure, much stronger than we would have predicted from our knowledge of current patterns. From the isolation by distance analyses, we saw that the scale of the population structure is much less than 500 km, which was the geographic scale on which the collections were sampled. In order to refine estimates of the geographic scale of POP population structure, we need to sample at a finer scale. Another conclusion we can make is that the geographic scale of the population structure appears to be less than the scale of the management areas (Figure 10). Finally, because of the strong structure observed in the adults, it is likely that most of young of the year fish that are displaced substantial distances by currents either perish or have the behavioral capability to return to their natal sites. Only a miniscule portion of juvenile fish survives to contribute to the adult population; and in many years there are virtually no survivals. Oceanographic influences clearly contribute to both the success and failure. However, those influences are not yet understood well enough to reliably be predictive.

Summary

Many marine species do not move long distances between birth and reproduction, even if environmental factors like strong oceanographic currents make it easy for them to move. One consequence is that the species is perpetuated at multiple sites, which produce young. Another is that if an area is overexploited, it may take a long time to recover because fish from nearby areas do not disperse and repopulate it rapidly. If the average age of reproducing females is 20 to 30 years (the generation time for Pacific ocean perch), a fishery in an overexploited area may be compromised for a much longer time (several generations).

The pattern of genetic variation that occurs over the geographic range inhabited by a species often reflects the geographic nature of its production. Is there a single large source of production, or are there multiple sources within the geographic range? Here we used the genetic compositions of collections of adult Pacific ocean perch sampled from multiple locations (neighborhoods) in the Gulf of Alaska to characterize its population structure. We learned that even though strong currents move counterclockwise around the Gulf of Alaska, a genetic structure exists at a spatial scale much smaller than the 500 km sampling scale. That scale is also smaller than the management scale. A parallel study of young-of-the-year Pacific ocean perch generally showed a similar structure, but there were indications that prevailing currents as well as physical features influenced their distributions. The strong structure observed in the adults, however, suggests that juvenile fish that are swept too far from “home” probably perish.

8/2014