A Primer of Evolution

An Introduction to Evolutionary Thought: Theory, Evidence, and Practice

Michi Tobler

Updated on 2022-07-19

Contents

Preface

In the Light of Evolution

> "There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been and are being evolved."
>
> Darwin, 1859

The concluding sentence of Charles Darwin's paradigm-shifting book "The Origin of Species" has stuck with me ever since I first read it in my freshman year of college. Back then, I was passionate about two things: the wildflowers of the Alps in my home country of Switzerland, and the fishes I kept in my aquaria. They both offered an incredible diversity of forms to explore. On the one hand, there were bell flowers (*Campanula* sp.; Figure **??**) and gentians (*Gentiana* sp.) with exuberantly large and colorful flowers, snowbells (*Soldanella* sp.) that pushed their flowers through a cover of snow, and willow (*Salix* sp.) brushes that barely managed to grow ankle-high in the many decades of exposure to the harsh mountain climates. On the other hand, there were the cichlid fishes of Lake Malawi. Over 1,000 species in a single lake, all carrying their developing young in their mouths but otherwise so different in their body morphology, coloration, and feeding habits. Some species specialize on scraping off scales from the bodies of other fish, others adapted to prey exclusively on fish fry, ramming mouth-brooding females in the throat and gobbling up their offspring as they are released. It's so wild that it seemed made up! I could get lost in exploring all the magnificent and weird things in nature. I still do sometimes; whether it's by stomping around creeks, observing my fish, or browsing through books. And all along I have kept wondering: Why are there all of these forms? How do these critters function in their habitats? And how did they come about?

Studying evolutionary biology has helped me to address these questions. As a scientific discipline, evolutionary biology fundamentally seeks to understand biodiversity and its origins. Evolution is the unifying theory of biology because

Figure 1: *Campanula scheuchzeri* (Scheuchzer's Bellflower) in the Swiss Alps. Photo: M. Tobler

it provides a simple explanation for the patterns of similarities and differences we can observe among all living things, which ultimately forms the framework in which researchers across disciplines address questions about the living world. It does so by addressing both proximate and ultimate causes of organismal function. Proximate questions primarily focus on explaining organismal function in terms of intrinsic and environmental factors (these are sometimes referred to *how* questions). For example, when we consider a peacocks magnificent tail feathers and coloration, how did environmental cues and changes in hormone levels initiate the development of these secondary sexual traits, and what genes might be involved in controlling there expression? In contrast, ultimate questions explain organismal function in terms of the evolutionary forces acting on them (sometimes referred to as why questions): why did the peacock evolve its exuberant traits, what are the traits' functions and how do they impact the fitness of its carrier?

Understanding how organisms work, how their traits allow them to survive in the peculiar environments they find themselves in, and how those traits came to be did not take out the wonder out of my fascination with nature; rather, it filled me with a new level of appreciation for the intricacies of life. I think that is *the grandeur* Darwin was alluding to at the end of *The Origin of Species*. Or, as another famous evolutionary biologist put it:

> "Nothing in biology makes sense except in the light of evolution. [...] Seen in the light of evolution, biology is, perhaps, intellectually the most satisfying and inspiring science. Without that light it becomes a pile of sundry facts, some of them interesting or curious, but making no meaningful picture as a whole."
>
> Dobzhansky, 1973

Beyond Understanding Biodiversity

Simply put, evolution is the change in heritable traits of populations that occurs across successive generations. Today, we have a nuanced understanding of the mechanisms that contribute to the evolutionary process. We are disentangling the genomic basis of traits relevant for organismal function, we are identifying the evolutionary forces—like natural selection—that determine what traits are passed on from one generation to the next, and we try to link these mechanisms to observable evolutionary outcomes, like adaptation, cooperation, and speciation. Applying these approaches has allowed us to explore many aspects of biodiversity, including the reasons behind sexual dimorphism, puzzling social behaviors, variation in life history traits, and even our own human origins.

The scientific reach of evolutionary biology, however, has long since eclipsed a basic understanding of the origins and function of biodiversity; the power of evolutionary analyses is now applied to address some of the major scientific challenges we face as a society: How will nature respond to the rapid environmental changes caused by human activities? How do we safeguard food production for a rapidly growing population? Why are there cancers and other diseases, and what can their origins tell us about prevention and treatment? How do we limit the spread of antibiotic-resistant pathogens? And how can we predict and limit the spread of emerging infectious diseases?

During this semester, we will cover both basic and applied aspects of evolution. So whether you are a bit of a naturalist—like myself—who tries to better understand the world, or whether you aspire to address some of the major environmental and public health issues we face as a society, I hope you will find something to take away from this class.

An Overview of the Semester

The semester, and accordingly this book, is structured into four parts, each with multiple chapters that build on each other. Each chapter corresponds to a weekly module.

Part 1: The Basics

In the first part of our journey, we will establish the basics of evolutionary biology. **Chapter 1** introduces the concept of evolution and provides some historical context of how Darwin conceived his idea of "descent with modification" to describe the pattern of evolution. Treating the idea of evolution as a hypothesis, we will also develop testable predictions that can be verified through observation or experimentation. **Chapter 2** will take a closer look at those predictions, and we will explore different lines of evidence that evolution has been—and still is—happening. Finally, **Chapter 3** will introduce Darwin's other big idea, natural selection, which describes a mechanism that can account for the tendency of organisms having traits well suited to the environments they live in.

Part 2: A Genetic Perspective on Evolution

The second part takes a 21st century perspective on evolution and closes a critical gap in Darwin's original ideas—namely, the mechanisms underlying heredity. We will integrate your knowledge of modern genetics with evolutionary principles to analyze changes in the genetic composition of populations through time. **Chapter 4** will explore how different types of mutations impact the expression of phenotypic traits and provide the raw material for evolutionary change. In addition, we will learn how evolutionary biologists quantify genetic variation in populations and use that data to infer whether or not evolutionary forces are acting on a population. In **Chapter 5** and **Chapter 6**, we will integrate evolution and genetics and use mathematical models to explore how natural selection interacts with other evolutionary forces (mutation, genetic drift, and migration) to shape the genetic composition of populations. In **Chapter 7**, we will investigate the evolution of DNA sequences, explore the molecular signatures of selection, and see what we can uncover about historical processes simply by interpreting patterns of DNA sequence variation. And finally, in **Chapter 8**, will learn about basic quantitative genetic approaches used to study the evolution of complex phenotypic traits controlled by many genes at once.

Part 3: Evolutionary Outcomes in the Real World

The third part of the book explores the outcomes of natural selection and other evolutionary forces. **Chapter 9** focuses on how we can infer adaptation in natural populations. We will explore the concept of phenotypic plasticity, why it can complicate the inference of adaptation, and how plasticity itself can be the outcome of adaptive evolution. In **Chapter 10**, we will focus on how evolution has shaped the social interactions between individuals of the same species, learning about kin and sexual selection. Finally, **Chapter 11** investigates how new species arise. We will discuss speciation as a gradual process that is shaped

by the same evolutionary forces that influence the evolution of phenotypic traits within species.

Part 4: Applied Evolutionary Biology

The last part of the book focuses on the application of evolutionary theory in the context of human nature and human health. **Chapters 12** and **Chapter 13** explore how evolutionary principles are applied in modern medicine. We will discuss why we age and how modern lifestyles are connected to the development of a wide variety of health conditions commonly named "diseases of civilization". In addition, we will explore how a better understanding of pathogen evolution allows for the development of concrete management strategies that can impact the spread of diseases. Finally, **Chapter 14** includes an overview of human origins and discusses how the sequencing of ancient DNA has shed new light into our own history.

How to Use This Book

This book is not designed to provide a comprehensive overview of current evolutionary biology. Rather, it is supposed to provide a succinct introduction to evolutionary thought revolving around theory, evidence, and practice:

- I will introduce some of the **theoretical cornerstones and core concepts** of modern evolutionary biology. The goal is for you to be able to apply these concepts and articulate testable hypotheses that explain natural phenomena from an evolutionary perspective.
- You will become familiar with the diversity of **empirical approaches and lines of evidence** that scientists use to address evolutionary hypotheses.
- You will practice approaching problems like scientists and evaluate data to address evolutionary hypotheses. To do so, you will learn how to program in R to **analyze and visualize data** and articulate your interpretations and conclusions.

In accordance with these goals, each chapter will provide you with a conceptual introduction to the topic. There is not much emphasis on examples, as we will explore those together in our class meetings. Each chapter also includes additional resources that you can explore if you have difficulties understanding or if you want to explore a topic in more detail. To help you with the R exercises, each chapter also provides you with additional background on case studies and R programming tutorials that help you to develop the necessary skills. Finally, each chapter ends with a series of reflection questions that will prompt you to review your learning. For quick reference, Appendix A summarizes some of the

key R code you will work with in the class, and Appendix B offers all exercises for quick download.

Please note that this resource has been developed and optimized to be used as an HTML book that can be accessed with any web browser. Accessing the book this way allows you to make full use of the dynamical content and R components, which are more limited if you print. Plus, you can save a bunch of trees:)

References

- Darwin, C (1859): *On the origin of species based on natural selection, or the preservation of favoured races in the struggle of life.* London: John Murray.
- Dobzhansky, T (1973): Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher* 35: 125–129.

Front Matters

Acknowledgments

I alternate teaching K-State's Evolution course (BIOL 520) with Tom Platt, and over the years, John Coffin, Ryan Greenway, Madison Nobrega, and Libby Wilson have served as graduate teaching assistants. They have all made immeasurable contributions to the development of the learning materials included here. I especially thank Madison Nobrega for her relentless editing; her input has greatly improved this text. Emily G. Finch from the Center for the Advancement of Digital Scholarship at the Kansas State University Libraries and Greg Dressman, Director of Enterprise Server Technologies at Kansas State University, provided critical logistical support.

This book was made possible with the support of a grant from the The Open/Alternative Textbook Initiative at Kansas State University.

Last but not least, many thanks to the past students in BIOL 520 at K-State. It was your input that lead to this resource and that helps me to continuously improve it. Especially, the following people have graciously provided feedback on writing and content: Shaun Baughman, Caroline Gatschet, Aaron George, Kate Odgers, and Andy Su.

Copyright

© 2022 by Michi Tobler

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. You are free to copy and redistribute the material in any medium or format, and to remix, transform, and build upon the material provided you give the original author credit (attribution), you

are not using the material for commercial purposes (NonCommerical), and you distribute your work under the same license as the original (ShareAlike).

If you would like to adopt and use this resource in any way, feel free to contact Michi Tobler to obtain a copy of all source files.

About the Author

Michi Tobler is a Professor in the Division of Biology at Kansas State University in Manhattan, KS. His research focuses on mechanisms of adaptation and speciation, primarily using extremophile fishes as study systems. You can learn more about our research on his lab website: https://sulfide-life.info/

Part I

Basic Concepts

Chapter 1

What Evolution Is

The notion of an ever-changing world is a relatively new idea in Western cultures. Dating back to the teachings of Greek philosophers, the world was long seen as permanent and unchanging. Its inhabitants were thought to be arranged on scale from lower to higher beings by a divine creator, with humans at the very top of the hierarchy. Until the early 1800s, most naturalists were primarily concerned with describing and cataloging life, which led to the modern classification system originally developed by the Swedish biologist Carolus Linnaeus.

By the time Charles Darwin was born in the early 19th century, evidence for a changing world was accumulating. Geologists started to contemplate how slow, gradual processes could carve canyons into stone and change the course of rivers. The discovery of fossils indicated the past existence of organisms vastly different than today's and suggested a rich history with continuous change. And some scholars, like the French biologist Jean-Baptiste Lamarck (Figure **??**), started to formulate theories that could potentially explain the diversity of life. Budding evolutionary thought was not limited to Western cultures, and there were a number of pre-Darwinian Muslim scholars who articulated evolutionary ideas as well (Malik et al. 2018).

It was Lamarck, inspired by the similarities he observed across different species, who formulated the first theory of evolution. Lamarck argued that life was inevitably driven towards increasing complexity, essentially progressing along the scale from lower to higher beings established by the Greeks. He envisioned that simpler life forms, like microbes, continuously arise and eventually develop into higher forms, like plants and animals. Lamarck also believed that progress along this hierarchy allowed organisms to adapt to their environment, which was driven by an "inner need". He thought that the continuous use of a particular organ allowed for its proliferation and subsequent inheritance to the next generation. Perhaps most famously, Lamarck explained that the long neck of the giraffe was the product of a continuous stretching toward leaves high on the

Figure 1.1: Jean-Baptiste de Monet Chevalier de Lamarck. Painted by Charles Thévenin, [Public Domain](https://creativecommons.org/share-your-work/public-domain/).

trees. So, across many generations of stretching, giraffe necks gradually reached their current size.

1.1 Darwin and the Conception of a New Idea

When Darwin boarded the British Navy Ship HMS Beagle as a 22-year old, he was already well trained in geology, chemistry, and natural history. Darwin was hired as an unofficial naturalist and companion to the captain of the ship, Robert FitzRoy, and the five-year journey around the world allowed young Darwin to collect evidence of a changing world (Figure **??**).

Darwin's visit to the Galapagos Islands, and the specimens of reptiles and birds that he collected there, proved to be particularly inspiring. The Galapagos are an archipelago of 21 small islands, located over 900 kilometers off the coast of Ecuador, and situated right on the equator. Darwin not only noticed that the fauna of these islands was dominated by animals found nowhere else, but there also were differences in the species from one island to the next. Most importantly, bird specimens that Darwin believed to be blackbirds, warblers, wrens, and finches due to their different beak morphologies later all turned out to be just finches—the Darwin's finches we know today (Figure **??**). In his own words:

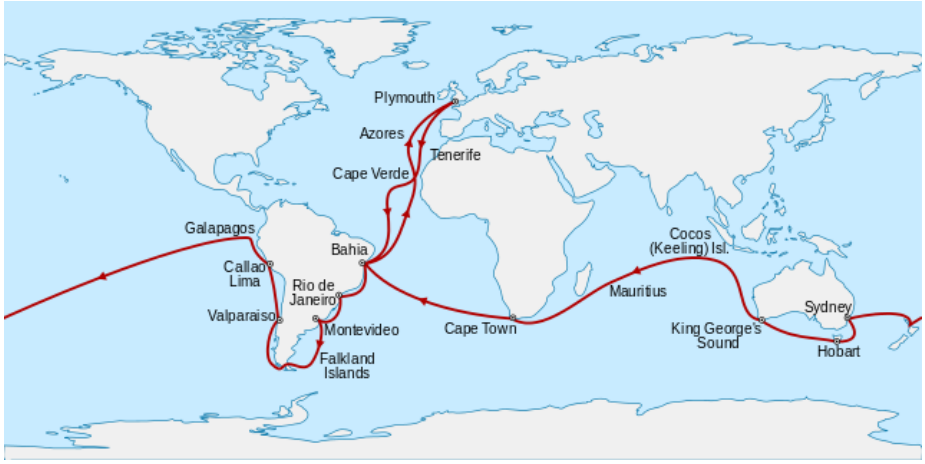


Figure 1.2: Map of the Voyage of the Beagle. Map: Sémhur, [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/).

> "I have stated, that in the thirteen species of ground-finches, a nearly perfect gradation may be traced, from a beak extraordinarily thick, to one so fine, that it may be compared to that of a warbler. [...] Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that from an original paucity of birds in this archipelago, one species had been taken and modified for different ends."
>
> — Darwin, 1889

If all the species were created with traits that fit the environments in which they are now found, why would the finches of the Galapagos Islands be so different from one another? Why do closely related species occupy niches that are filled by very different taxa in other regions? Observations like these, which Darwin made at many destinations throughout his travels, led him to suspect that the species we observe today have evolved from a shared ancestor.

Upon his return to England in 1836, Darwin settled near London and never traveled abroad again. Nonetheless, he remained a prolific scholar writing monographs about various topics, from geology to barnacles. Darwin was also a passionate pigeon breeder, and it was this unlikely hobby that provided additional insights for the formulation of his ideas. It led him to think about the forces that could drive the modification of species from one common ancestor to different descendants:

> "It is, therefore, of the highest importance to gain a clear insight into the means of modification. [...] At the commencement of my

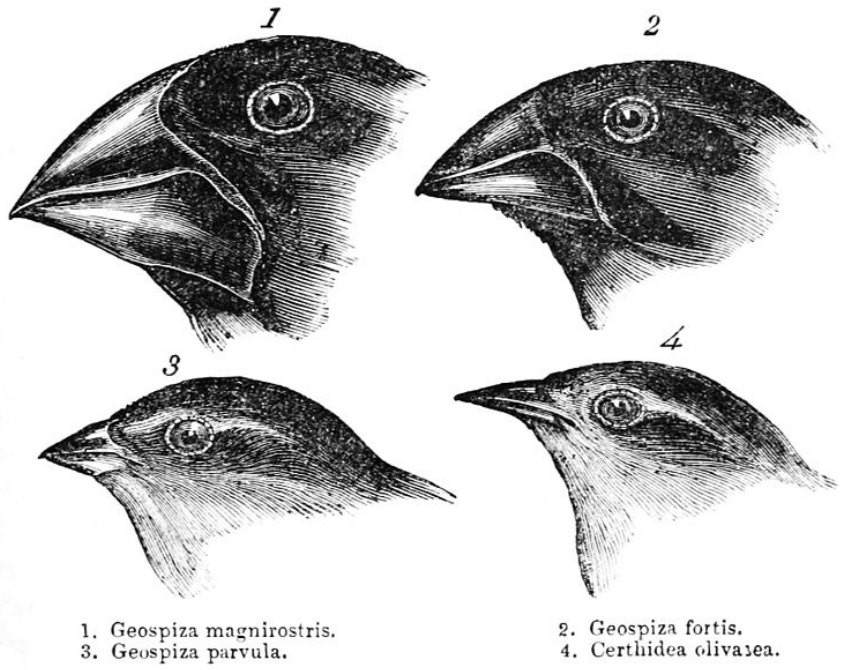


Figure 1.3: Examples of Darwin's finches or Galapagos finches. Note the stark variation in beak sizes among the four species. Drawn by John Gould, [Public Domain](https://creativecommons.org/share-your-work/public-domain/).

> observations it seemed to me probable that a careful study of domesticated animals and of cultivated plants would offer the best chance of making out this obscure problem."
>
> — Darwin, 1859

In his endeavors of pigeon breeding, Darwin noticed two critical things: (1) Offspring tended to inherit the traits of their parents. (2) If he carefully selected breeders with desired traits generation after generation, he was able to shape the variation of colors, morphologies, and behaviors in his flock. It was his meticulous work as a natural historian, collecting evidence for his emerging ideas for over two decades, that ultimately allowed Darwin to formulate the basic tenets of evolutionary biology that still hold up today, after more than 150 years of scrutiny by the scientific community.

1.2 Two Fundamental Insights

1.2.1 The Pattern: Evolution is Descent with Modification

The first of Darwin's fundamental insights was the formal description of evolution as we understand it today. He described evolution as *descent with modification*, postulating the common ancestry of all living things. Hence, different species did not arise independently, but they derived from preexisting form. This insight arose from a careful analysis of patterns of similarities across species. Species that share a recent common ancestor share traits precisely because they inherited them the shared ancestor. More distantly related taxa exhibit differences in their traits, because they have been on independent evolutionary trajectories for longer periods of time. This perspective describes the observable *pattern of evolution*.

Given Darwin's insight, here is a simple definition of evolution that we will rely on for now. As you will see, there are multiple modern definitions of evolution, and we will revisit different definitions throughout the book.

Definition: Evolution

Evolution is the change in the inherited traits of a population across successive generations, ultimately leading to the transformation of species through time (both in terms of changes of traits that occur within species and the origin of new species).

The power of Darwin's idea of descent with modification is that we can treat it as a scientific hypothesis with empirically testable predictions. If Darwin's notion of evolution was right, we should be able to uncover evidence that:

1. Species change through time (microevolution).

2. Lineages split to form new species (speciation).
3. Novel forms derive from earlier forms (macroevolution).
4. Species are not independent but connected by descent from a common ancestor (common ancestry and homology).
5. Earth and life on Earth are old (deep time).

We will revisit these predictions and examine the evidence for evolution in detail in Chapter 2.

1.2.2 The Process: Natural Selection is a Mechanism of Evolution

The second of Darwin's fundamental insights was the inception of a mechanism that could produce the observable pattern we call evolution (i.e., change in inherited traits across generations), which Darwin named natural selection.

Definition: Natural Selection

Natural selection is the process in which individuals with a particular trait exhibit higher reproductive success than individuals without that trait.

Natural selection explains how the traits of a population change through time, and why organisms are well suited for their environment. Individuals that exhibit traits that are advantageous under certain environmental conditions have a higher chance of surviving and reproducing, making a disproportional contribution to the offspring of the next generation. If the relevant traits are heritable, their frequency increases across subsequent generations. This process not only leads to change in populations across generations, but said change specifically pertains to traits that are important for survival and reproduction in a given environmental context. In other words, the action of natural selection directly leads to adaptation. We will take a close look at how natural selection works in Chapter 3.

1.2.3 Distinguishing between Pattern and Process

A key misconception is that evolution and natural selection are the same. People often use the two terms interchangeably, conflating patterns and processes of evolution. Distinguishing between the two is critical, because evolution and natural selection do not have to go in unison.

Evolution is a historical pattern of change that can—but does not have to—be caused by natural selection. Evolutionary change can also be driven by other forces that impact the composition of populations across generations. In Part II of this book, we will examine evolutionary forces other than natural selection, which include mutation, genetic drift, and migration. In natural populations, these four forces interact to shape evolutionary change across time.

It is also important to note that the action of natural selection does not necessarily lead to evolution. Natural selection can only cause evolution when it acts on heritable traits that are transmitted from parent to offspring. If that is the case, the offspring of successful individuals will carry the same traits that made their parents successful. However, not all traits are heritable. For example, selection on individuals exhibiting high muscle mass does not translate to evolutionary change if that muscle mass was acquired through exercise and diet. If natural selection acts on non-heritable traits, there is no evolutionary change.

Explore More: Misconceptions about Evolution

Despite the broad scientific consensus on the importance of evolutionary theory in understanding the diversity of life, people hold many misconceptions about evolutionary theory and processes, its implications, and its relation to religious beliefs. I encourage you to explore the fantastic resource provided by The University of California Museum of Paleontology, which lists common misconceptions about evolution and clarifies these misconceptions.

1.2.4 Population Thinking

Darwin's fundamental insights were possible because of a fundamental shift in how he thought about biological entities, from a typological perspective to what Mayr (1982) called "population thinking". Prior to Darwin, scholars viewed individual organisms as imperfect representations of a central Platonic type (the perfect manifestation of a species). Variation among individuals was considered to be meaningless noise that was either ignored or seen as a nuisance when describing and classifying species. In contrast, population thinking rejects the notion of an ideal representative and instead focus on the variants found within populations. Variation among individuals is not just meaningless noise that represents different degrees of imperfections from a preconceived idea, but it is the raw material for evolutionary change. In other words, what was meaningless noise to naturalists prior to Darwin suddenly became the very focus of evolutionary studies after. Mayr (1982) argued that this paradigm shift was one of Darwin's most important contributions to modern biology.

1.3 Practical Skills: Getting Started with R

During this course, you will learn about evolution *and* practice being a scientist by visualizing and interpreting data on a weekly basis. To do so, we will use an open-source software called R. This section provides instructions for installing R and a companion program (RStudio) on your Mac or PC.

1.3.1 What are R and RStudio?

R is a programming language for statistical analysis and data visualization, and over the past decades, it has become one of the most critical and universally used tools in the life sciences. RStudio is a convenient interface that allows you to easily write, organize, and execute your R code; it's the program we will use throughout the semester for weekly assignments. Note that RStudio will not run unless R is already installed.

1.3.2 Installing R

To run R and RStudio on your system, complete the following steps in the right order (*i.e.*, you need to install R before you can install RStudio). Although we recommend the use of Windows or Mac OS for this class, R is also available for Linux, and it can be installed on Google Chromebooks. Please note that any code used in this course has not been troubleshot on the latter two platforms.

For Windows

- Download the binary setup file for R 4.2.x
- Open the downloaded *.exe file and follow the instructions to install R
- Download the binary setup file for Rtools40
- Open the downloaded *.exe file and follow the instructions to install Rtools

For Mac

- Check the version on your operating system (click the Apple logo on your desktop and choose "About This Mac"). If your macOS version is 10.14 or older, please first update your computer to a newer macOS version. If you do not update your computer first, R and RStudio may still install properly, but you will run into compatibility issues later in the semester when use different packages within R.
- Download the .pkg file for R 4.2.x form the following link: https://cran.r-project.org/bin/macosx/
- Open the downloaded *.pkg file and follow the instructions to install R

1.3.3 Installing RStudio

Download the free, open-source version of RStudio Desktop by choosing the appropriate installer file for your operating system and then run it to install RStudio. If you are using Mac, make sure to move the RStudio app to your actual Applications folder (rather than starting it from the disk image).

1.3.4 Checking the Successful Installation

Once you completed the steps above, we recommend that you launch RStudio and make sure it starts without any errors. If the installation was successful, the RStudio interface should look something like Figure **??**:

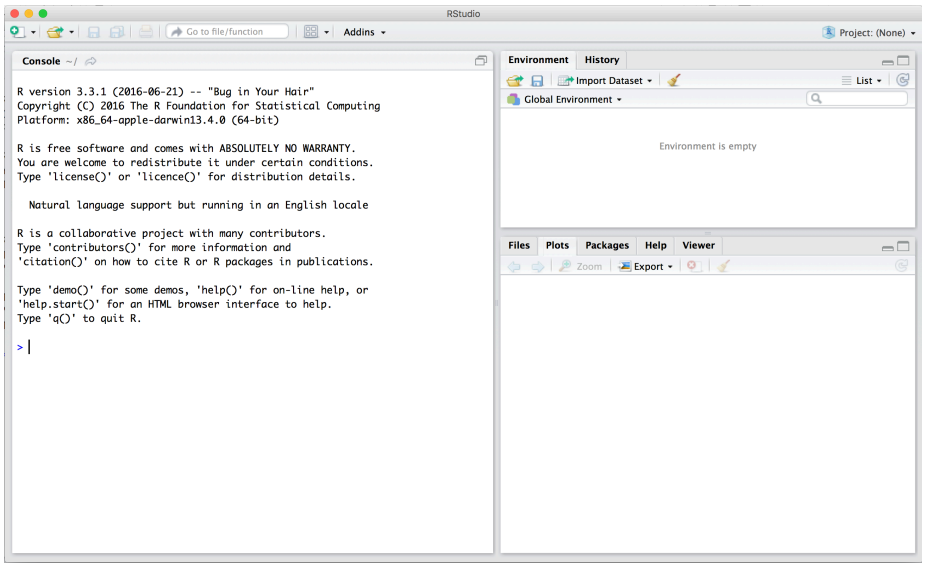


Figure 1.4: Screenshot of the RStudio interface.

1.3.5 Some R and RStudio Basics

This section was inspired by and used parts of "Introduction to R" by Russell A. Poldrack, published on LibreTexts and licensed under CC BY-NC.

Learning a programming language—like learning any language—can be intimidating. Still, it is a basic skill for any scientist. Starting in Chapter 2, you will work through R exercises that help you digest the basic concepts discussed in each chapter. We will show you exactly how to use RStudio to work through the exercises there. As you will see, most assignments in class will provide you with the majority of the code structure in a way that a few simple tweaks on

your side will produce the desired outcome. But we encourage you to explore beyond what the exercises dictate. A willingness to experiment will make you a better programmer and scientist. Before we get to that stage, however, you need to build some basic fluency in R, so you can ease yourself into using this new tool.

The Console and Basic R Prompts

Let's begin with the fundamental building blocks of R and RStudio. As you open RStudio, the program window will look something like Figure 1.4.The panel in the upper right contains your work space (Global Environment) as well as a history of the commands that you've previously entered. Any plots that you generate will show up in the panel in the lower right corner and you can also use that section to browse your folders for specific files. The panel on the left is where the action happens. It's called the Console. Every time you launch RStudio, it will have the same text at the top of the console telling you the version of R that you're running. Below that information is the command prompt (symbolized by `>`), where you can enter commands and R responds to those commands.

It's really simple, and you can try it out right now! Just type `2+2`, press enter, and R should respond right back with the response. It should look like this:

```
2+2
```

```
## [1] 4
```

R pretty much operates like a calculator and can perform any calculation you might need. R can also deal with logical variables, and respond to queries about whether logic statements are true or false. For example, you can use logical operators [`>` (greater than), `<` (smaller than), `==` (equals), `!=` (not-equals)] to contrast statements, and R will assess them:

```
5*2>sqrt(100) #note that sqrt starts for square-root
```

```
## [1] FALSE
```

```
5*2==sqrt(100)
```

```
## [1] TRUE
```

Working with Variables

One of the strengths of R is that it cannot only perform basic calculations using numbers, but it can handle variables, which are symbols that stand for values. You can designate a variable using `<-`. For example, we can assign the value 5 to x:

```
x <- 5
```

Notice when you run that code, you created a new variable x with the value 5 that shows up in your work space on the top-right panel. Once you have a variable defined, you can use it in calculations and logic statements as described above:

```
x^2
```

```
## [1] 25
```

```
x==5
```

```
## [1] TRUE
```

Designating variables allows you to write some complex code that processes data, and if you want to do the same thing with different data, you don't have to rewrite the code, you just redefine the input variables. For example you can designate three different variables:

```
x1 = 16
x2 = 12
x3 = 2
```

If you use these variables in a complex equation, you need to just write the equation once, and you can tweak the input variables to explore the effects of different values.

```
x1*(x1+x2+x3)/(x1^2+x2^2+x3^2)
```

```
## [1] 1.188119
```

Vectors

Variables are not only able to designate numbers, they can also designate lists of numbers (called vectors). You can create a vector (`y` in this case) using the `c()` function, and you can then call on `y` using `print()` for R to display the vector in the console:

```
y <- c(1,12,2,14)
print(y)
```

```
## [1]  1 12  2 14
```

If you execute this code, you will again see the creation of a new variable y in your work space, which corresponds to a vector with four numbers.

You can call numbers within a particular vector by using square brackets along with the variable name and a the number that refers to a location within the vectors. So, if we want to extract the 3rd number in our vector `y`, we simply type:

```
y[3]
```

```
## [1] 2
```

We can also extract a range of values by using a colon. Let's say you want to extract the last two numbers in our vector:

```
y[3:4]
```

```
## [1]  2 14
```

Finally, you can modify specific values within a vector by just combining some of the code you have already learned so far. For example, you can change the second value in our vector `y` from a 12 to 4:

```
y[2] <- 4
y
```

```
## [1]  1  4  2 14
```

Using Functions

R can execute a wide variety of functions that take some input and provide an output based on the input. Functions can be applied to a simple number. For example the `sqrt()` function calculates the square-root of an input number. We can use it on our variable x that we defined above:

```
sqrt(x)
```

```
## [1] 2.236068
```

We can also apply the same function to a vector, in which case the function is applied to every number in the vector:

```
sqrt(y)
```

```
## [1] 1.000000 2.000000 1.414214 3.741657
```

Sometimes we want to use the output of one function as an input for another function. In this case, it makes sense to store the output of a function in a new object:

```
sqrt.y <- sqrt(y)
print(sqrt.y)
```

```
## [1] 1.000000 2.000000 1.414214 3.741657
```

Note that this code again created a new object called `sqrt.y` in your work space, and you can call on it using `print()` to display it in the console. We can now also use the new values as input for another function. For example, we can calculate the average of these values using the `mean()` function:

```
mean(sqrt.y)
```

```
## [1] 2.038968
```

Over the course of the semester, you will get to know a wide variety of functions in R that will allow you to make complex plots and analyze genetic and phenotypic data to make evolutionary inferences. R already knows a large number of functions upon installation, but you can essentially teach it limitless new functions by installing so-called libraries (or packages). There are thousands of libraries for R, and they contain functions to analyze and visualize all kinds of biological data, from the structure of genomes to the composition of ecological communities. You will learn how to install new libraries and apply the functions they contain in Chapter 2.

Data Frames

The reality in biology is that we rarely deal with simple numbers or vectors. Rather, we typically collect complex data sets that contain a multitude of variables. For example, if we go out into natural populations to quantify trait variation, you might record information about an individual's sex, its health status, body size, and body mass. Instead of having a different variable designated for each of these pieces of information, we can combine all of the variables into a single object (table) called a data frame.

To assemble a data frame, we start by first defining individual variables. In the example below, I create two vectors with categorical variables (`sex` and `health.status`; note that different categories need to be contained in quotation marks), and two continuous variables (`body.size` and `body.mass`):

```
sex <- c("male", "female", "male", "female")
health.status <- c("healthy", "healthy", "sick", "sick")
body.size <- c(16.4, 12.2, 10.2, 8.9)
body.mass <- c(221, 199, 178, 159)
```

As before, executing this code will generate four different vectors in your work space. To combine all the information into a single data frame, we can use R's `data.frame()` function. Within the data frame, you can generate a variable name (before `=`) and assign a vector that you already created (after `=`):

```
df <- data.frame(sex=sex, health=health.status, size=body.size, mass=body.mass)
```

Executing this code will again generate a new object in your work space. You can view your data frame by double-clicking on it in the workspace or using the `View()` function in the console:

```
View(df)
```

The data frame should look similar to a spreadsheet you might know from Excel (Figure **??**):

	sex	health	size	mass
1	male	healthy	16.4	221
2	female	healthy	12.2	199
3	male	sick	10.2	178
4	female	sick	8.9	159

Figure 1.5: A view of the df data frame generated by the 'View()' function.

Each of the columns in the data frame contains one of the variables, with the name that we gave it when we created the data frame. We can access each of those columns using the `$` operator. For example, if we wanted to access the size variable, we would combine the name of the data frame with the name of the variable as follows:

```
df$size
```

```
## [1] 16.4 12.2 10.2  8.9
```

This is just like any other vector, in that we can refer to its individual values using square brackets, as we did with regular vectors:

```
df$size[3]
```

```
## [1] 10.2
```

In reality, you will rarely have to build your own data frame in this class, because we will provide you will real data from classic studies and our own research that you can read into R. You will learn how to do that in Chapter 2.

Help Functions

If you ever run into issues or have question regarding how to use a particular function or package, you can access the documentation with instruction using the question mark symbol (`?`) followed by the function or package name. For example, `?sqrt` will provide the instructions for the `sqrt()` function.

1.4 Additional Resources

1.4.1 Alternative Evolution Textbooks

If you would like more in-depth reading materials about evolution, there are a number of excellent textbooks available. I particularly recommend the edited volume by Losos et al. with short articles about current topics in evolution, because it is freely available for online reading through most university libraries.

- Bergstrom CT, LA Dugatkin (2018): *Evolution*. Norton.
- Emlen DJ, C Zimmer (2020): *Evolution - Making Sense of Life*. MacMillan.

- Futuyma DJ, M Kirkpatrick (2017): *Evolution*. Oxford University Press.
- Herron JC, S Freeman (2014): *Evolutionary Analysis*. Pearson.
- Losos JB, DA Baum, DJ Futuyma, HE Hoekstra, RE Lenski, AJ Moore, CL Peichel, D Schluter, MC Whitlock (editors) (2014): *The Princeton Guide to Evolution*. Princeton University Press.

1.4.2 Evolution in the Primary Literature and the News

For some assignments, and to satisfy your innate curiosity, you may want to consider consulting other resources. You can browse through the list of peer-reviewed journals that publish research related to evolution. If you want to find current research in the field, I also recommend the following websites, many of which provide lay summaries of recent papers:

- EurekAlert!
- Nature News
- Popular Science
- Science Daily
- Science Magazine News
- Science News
- Scientific American

1.4.3 R and RStudio Resources

This book will provide you with the background knowledge necessary to use R at the level you need to succeed in this class. Nonetheless, you will run into issues, and being able to troubleshoot errors is one of the most important coding skills. **Perhaps the most important tool during troubleshooting is... Google!** If you google your error message from R or specific questions you have, you will likely find that somebody else already had that problem. Question/answer threads associated with Stack Overflow usually are a very reliable resource for overcoming issues with coding. If you need additional resources or want to dig a little deeper, consider the following:

- Getting Started with R: An Introduction for Biologists (this may be worth it if you consider attending graduate school to earn a M.S. or Ph. D.)

- The Statistical Tools for High-Throughput Data Analysis website has a number of fantastic R tutorials. The following may be helpful for you in this course:
  - Basic introduction to R and RStudio
  - Import data into R
  - How to make any ggplot
  - Introduction to ggplot2
  - Bar plots with ggplot2
  - Histogram plots with ggplot2
  - Scatter plots with ggplot2
  - Box plots with ggplot2

1.5 Reflection Questions

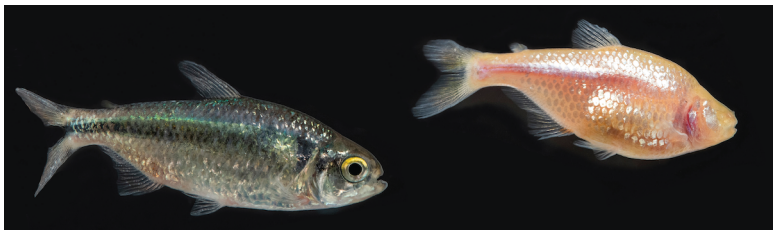
1. How would you define evolution in your own words? How does your definition compare to the one given in this chapter? Are there other definitions of evolution? What are their strengths and weaknesses?
2. What is the difference between evolution and natural selection?
3. What do you think of the term "survival of the fittest"? Is it an accurate description of evolution? If not, why?
4. Do you think evolution is a fact or a theory?
5. Darwin once wrote (1859): "We see nothing of these slow changes in progress, until the hand of time has marked the lapse of ages." How long do you think it takes for evolution to take place?
6. While Darwin understood that some traits are inherited from parents to their offspring, he did not know exactly how that happens. He was largely unaware of the work of his contemporary, Gregor Mendel, who worked out the foundational principles of genetics. Instead Darwin hypothesized that the body continuously emitted small particles he called gemmules that accumulated in the gonads and contributed heritable information to the gametes. How do you think Darwin's misconceptions about heredity impacted the validity of his evolutionary ideas? How have discoveries in classical genetics and later the discovery of DNA carrying the genetic instructions for the development and functioning of all organisms strengthened or weakened our understanding of evolution as conceived by Darwin?
7. How does Lamarck's view of evolution differ from our current understanding? What aspects did he get right, and what did he get wrong?

1.6 References

- Darwin, C. (1859). *On the origin of species based on natural selection, or the preservation of favoured races in the struggle of life.* John Murray.
- Darwin, C. (1868). *The Variation of Animals and Plants under Domestication.* John Murray.
- Darwin, C. (1889; original in 1839). *Journal of Researches Into the Natural History and Geology of the Countries Visited During the Voyage of H.M.S. "Beagle" Round the World, Under the Command of Capt. Fitz Roy.* Ward, Lock and Company.
- Malik, A. H., Ziermann, J. M., & Diogo, R. (2018). An untold story in biology: the historical continuity of evolutionary ideas of Muslim scholars from the 8th century to Darwin's time. *Journal of Biological Education*, *52*(1), 3–17.
- Mayr E. 1982. The Growth of Biological Thought: Diversity, Evolution, and Inheritance. Cambridge (MA): Harvard University Press.

Chapter 2

Evidence for Evolution



Darwin described evolution as descent with modification. It turns out that he was not the only one to think about the ever-changing world in this way. Another prominent naturalist of the time, Alfred Russel Wallace, independently conceived the theory of evolution through natural selection. Like Darwin, Wallace conducted extensive fieldwork in the tropics and was a meticulous observer of the natural world. In 1858, Wallace wrote a letter to Darwin—who was by then an eminent scholar but had not published his views on evolution yet—detailing his own ideas about natural selection. This led to the joint publication of short abstracts detailing Darwin's and Wallace's views of evolution, and more importantly, it motivated Darwin to finish and publish his famous work, *On the Origin of Species*, in 1859. So, why does most of the credit for formalizing evolutionary theory go to Darwin rather than Wallace? Well, Darwin was undoubtedly first, ruminating on his ideas about evolution for decades before deciding to publish. As a consequence, he was able to introduce his views in much richer detail and provided many lines of evidence in support of his theory.

Explore More

To learn more about Alfred Russel Wallace, listen to "*He Helped Discover Evolution, And Then Became Extinct*", an NPR story published on the 100th anniversary of his death.

So, what evidence do we have that evolution is actually happening? What is the evidence for the occurrence of change in inherited traits across successive

generations, the transformation of species through time, and the emergence of new species?

As mentioned in Chapter 1, we can (must!) treat Darwin's idea of descent with modification as every other scientific hypothesis and develop testable predictions that are falsifiable with data. The idea of descent with modification makes five predictions that we can address with data:

1. Species change through time (microevolution).
2. Lineages split to form new species (speciation).
3. Novel forms derived from earlier forms (macroevolution).
4. Species are not independent but connected by descent from a common ancestor (common ancestry and homology).
5. Earth and life on Earth are old (deep time).

This chapter takes a closer look at the different lines of evidence we have in support of evolution.

2.1 Microevolution

Microevolution is the change in inherited traits of a population from one generation to the next, ultimately leading to the accumulation of changes and the transformation of species through time. *Heritable trait* in this context can refer to any phenotypic trait (for example the average beak size in a population of a bird) or a molecular trait (for example the frequency of alternative alleles at a particular locus). While changes in most traits from one generation to the next are subtle at best, strong natural selection can lead to significant and detectable evolutionary changes in very short periods of time. For example, check out the following video produced by the Kishony Lab at Harvard Medical School. They have designed a simple way to observe how bacteria evolve as they encounter increasingly higher doses of an antibiotic and adapt to survive—and thrive—despite of it.

You might say that bacteria are different. After all, assuming a generation time of 30 minutes, the two-week experiment described in the video represents over 670 generations of bacterial evolution. Translated to humans, that would represent about 17,000 years. Looking back that far in history, that was a time when humans exclusively lived as hunter-gatherers and just started to migrate into North America over the Bering Land Bridge…

One of the most persistent misconceptions about evolution is that it takes millions of years to occur. However, the reality is that microevolution—in principle—can happen in as little as one generation. Those short-term changes can be very hard to detect, because our measurement error of a trait of interest is often larger than the actual per-generation evolutionary change. Nonetheless,

over the course of just a handful of generations, natural populations may exhibit significant evolutionary change that we can detect with high confidence using genetic markers (*i.e.*, measuring changes in allele frequencies) or phenotypic measurements.

The convergence of ecological and evolutionary timescales is a relatively recent insight. Darwin did not think that we would be able to directly observe evolutionary change over short periods of time:

> "We see nothing of these slow changes in progress, until the hand of time has marked the lapse of ages."
>
> — Darwin, 1859

However, with technological breakthroughs that improved the precision of measurements we take in natural populations and with scientists' ability to track populations continuously through time, we have accumulated data across dozens of study systems—from microbes to vertebrates—documenting microevolutionary change within a few to a few dozen generations (see Hairston et al. 2005; Carroll et al. 2007). Here, I will briefly introduce you to evidence for rapid evolution gathered in one such study system (the threespine stickleback). In this chapter's case study, you will explore another example based on a time series of beak size variation of a species of Darwin's finch.

2.1.1 The Case of Threespine Stickleback

Threespine stickleback (*Gasterosteus aculeatus*; Figure **??**) are a widely used system to study evolution and have been shown to rapidly adapt to novel environmental conditions. Stickleback are primarily marine and inhabit coastal waters throughout much of the Northern hemisphere. They are small fish (usually less than 8 cm in length) that exhibit exquisite adaptations to avoid predation in their environment: the sides of their body are covered in bony plates, and they have spines associated with their dorsal and pelvic fins that, when spread out, can dissuade a predator from capturing or consuming them.

Since the last ice age, as glaciers retreated and left behind a plethora of new streams and lakes, stickleback have also colonized freshwater habitats, which differ in many ways from the original marine habitats. Freshwaters not only exhibit a different water chemistry, but they also tend to harbor fewer predators and different food resources. Over the past 10,000-20,000 years, stickleback in freshwater environments have evolved a number of phenotypic differences compared to their marine ancestors, including a drastic reduction of the armor plates along the body and—in some instances—a loss of the pelvic spines (Jones et al. 2012). Moreover, stickleback have also adapted to different niches within freshwaters, and there are distinct morphs in streams and lakes, and in

Figure 2.1: Threespine stickleback (*Gasterosteus aculeatus*). Photo by Gilles San Martin, [CC BY-SA 2.0](https://creativecommons.org/licenses/by-sa/2.0), via Wikimedia Commons

benthic and pelagic habitats within lakes (Hendry et al. 2013). Different freshwater ecotypes exhibit distinct body shapes and colorations and are adapted to consuming different types of prey items.

So, how long might it take for the evolution of the traits that vary so drastically across different stickleback forms? Sure, 20,000 years is a blink of an eye in the history of life on the planet, but it is still an eternity for any researcher that might want to observe evolution in action.

One hint at how fast stickleback might evolve comes from a fascinating natural experiment. In 1964, the Great Alaska Earthquake brought widespread destruction to the region and literally reshaped the regional topology. For example, multiple islands in the Prince William Sound and the Gulf of Alaska were lifted up further out of the ocean, creating new freshwater ponds where previously were none. In the time since the earthquake, stickleback have colonized these new freshwater ponds, and within just 50 years, they have evolved similar phenotypic traits that we know from stickleback in continental freshwaters (Lescak et al. 2015). Hence, adaptation to freshwaters upon colonization from the ocean may occur in a matter of a few decades rather than gradually over thousands of years of evolution.

To get a better understanding of just how fast evolution may proceed, researchers from the University of Basel in Switzerland decided to conduct a field experiment using lake and stream stickleback (Laurentino et al. 2020). The researchers first sequenced the genomes of lake and stream stickleback to detect the genomic regions that are differentiated between ecotypes and likely contain the genes involved in shaping the phenotypic differences between them. After that, they generated F2 crosses between the ecotypes, which is like shuffling a deck of cards from a genomic perspective: individual F2 offspring es-

sentially exhibit a random mixture of genomic segments from their lake and stream ancestors. If these F2 hybrids were introduced into a stream environment, it is predicted that individuals that exhibit stream alleles in regions of the genome important for the expression of stream-specific phenotypic traits perform better than individuals with lake alleles. At a population level, this should lead to an increase in the frequency of alleles characteristic for stream stickleback. And, this is exactly what happened when the researchers actually conducted the experiment. More importantly, the predicted genetic changes were detectable within just one generation of F2 individuals being released into a stream habitat. So, when selection is strong and researchers have the capability to measure changes in traits with adequate precision, we can actually detect the small, generation-to-generation changes that ultimately accumulate to give rise to more conspicuous evolutionary changes that are easier to detect.

2.2 Speciation

Speciation is the process by which new species arise. Before we dive into how speciation actually works, we should agree on what species actually are:

Definition: Species

A *biological species* is a group of organisms that can reproduce with one another in nature and produce fertile offspring. Species are characterized by the fact that they are reproductively isolated from other such groups, which means that the organisms in one species are incapable of reproducing with organisms in another species. (As you may know, there are alternative definitions of a species, which we will discuss in more detail in Chapter 11.)

In phylogenetic trees, speciation is depicted as a singular point that represents the moment one lineage splits into two (red circle in Figure **??**). Although speciation can occur instantaneously, for example when polyploidization is involved (see Chapter 11), new species typically evolve gradually, from a single variable population, to populations within a species that are differentiated but still connected through gene flow, to distinct species that are completely isolated from each other (Figure **??**). Movement along this "speciation continuum" is driven by the accumulation of reproductive barriers that prevent individuals from mating or successfully producing offspring with each other. Importantly, movement along the speciation continuum can be bidirectional, and reproductive barriers can disappear such that two species merge back together into one. If speciation is a gradual process, we should be able to observe all stages along the speciation continuum in nature, not just the endpoints of the speciation process with reproductively isolated species.

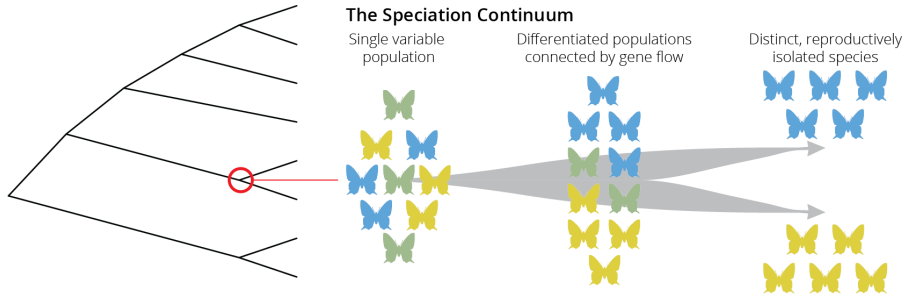


Figure 2.2: Speciation is not typically an instantaneous process. Rather species evolve gradually along a speciation continuum.

2.2.1 Ring Species

One phenomenon that perfectly illustrates the speciation continuum and provides evidence for ongoing speciation are so-called ring species, in which two reproductively isolated populations living sympatrically (red and brown in Figure **??**) are connected by a geographic ring of populations that can interbreed. Such ring species arise when an original population disperses around a geographic barrier, and populations diverge gradually, for example as a consequence of adaptation to local environmental conditions. Once populations come into sympatry again behind the geographic barrier, sufficient differences have accumulated such that populations cannot interbreed with each other anymore.

Several well-studied examples of ring species exist. For example, plethodontid salamanders of the *Ensatina eschscholtzii* complex have colonized different parts of California from the north and expanded southward around the Central Valley, which represents unsuitable habitat for salamanders. At the southern tip of the Central valley, salamander populations from the eastern and western mountain ridges that surround the valley came into secondary contact and are unable to interbreed due to the genetic changes that have accumulated during evolution in isolation (Pereira et al. 2011). Other examples of ring species include the herring and lesser black-backed gulls (genus *Larus*) that have a circumpolar distribution and cannot interbreed in northern Europe. In addition, greenish warblers (*Phylloscopus trochiloides*) form a ring species around the Himalayas (Irwin et al. 2005).

2.2.2 Catching Speciation in Action

Evidence for ongoing speciation also comes from a wide variety of study systems that do not occur in a ring. Especially when populations are subject to strong natural selection (for example if they inhabit different habitat types), we do not only observe adaptive differentiation across populations but also the inadvertent