How Do We Understand Natural Selection?

by

Lucy E Delaney B.S. (John Jay College of the City University of New York) 2012 M.A. (Hunter College of the City University of New York) 2016

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Sciences in the Graduate College of the University of Illinois at Chicago, 2021

Chicago, Illinois

Defense Committee: Joel Brown, Chair & Co-Advisor Boris Igić, Co-Advisor Roberta Mason-Gamer Alan Molumby Donald Wink, Chemistry ©Lucy E Delaney, 2021

For Dexter,

because of the implication.

#### ACKNOWLEDGMENTS

This thesis represents a mountain I was often not sure I could climb. That I am writing this today near the mountaintop is because of the love, energy, or support (sometimes all three) of a great many people. To begin, finishing a PhD requires first convincing someone to let you try. For that I am indebted to my advisor Joel Brown: for his endless enthusiasm during the twists and turns of my research these last six years, and for his unwavering belief in my success. This work starts and ends with that belief. My other advisor, Boris Igić, challenged me fiercely and held me to what often felt like impossibly and painfully high standards. Trying (and usually failing) to live up to those standards elevated my work in innumerable ways and produced a collection of "skills" that are the direct result of the pursuit of excellence at his behest. Both Joel and Boris will stay with me long after graduate school as the broader ecological and evolutionary worldviews that sit atop each of my shoulders. My committee members: Robie Mason-Gamer, who never let me wander too far into the weeds with my writing; Don Wink, who always reminded me to have empathy for our students in everything I do; and Alan Molumby, who made teaching introductory biology seem like the best job in the world.

Advisors and committees of course profoundly influence the grand shape and structure of a thesis, but peers influence the day-to-day grit of putting the thing together. I am eternally grateful to my best JBs, Andy and Karolis, for being there with me every day in the mud. This thesis exists because when I was there I was not alone. I love you both. The folks of the Brown Lab also never failed to brighten my spirits and consistently provided warmth and reassurance, even in my darkest days. Especially Meg, who always encouraged me to keep at least one foot on the ground.

#### ACKNOWLEDGMENTS (Continued)

For many years, I attended school as a commuter that also had a full-time job: I did not form a great many relationships with my Professors. One notable exception is my undergraduate philosophy advisor, James DiGiovanna. I reluctantly took his Ethics class as a required course during my first semester of college and it transformed the trajectory of my scholarly career. Stephen C. Stearns wrote, "To learn to think, you need two things: large blocks of time, and as much one-on-one interaction as you can get with someone who thinks more clearly than you do." He was (and still is) one of those people for me.

And for those that I love, who loved me without ever really understanding what I was doing: my students, the Bouzek-Cross family, the Cabey family, the Kreston-Brody family, the Layton family, the Long-Davidson-Singer-Stern family, the Lenhart-Weis family, the Mandolini-Trivino family, the Quane family, and the Sadler family. This love brought me from the depths again and again. Nate, for being there cheering in the parents pit the day I graduated from college and for being there cheering as I finish my PhD a decade later. My family—Stan, Margaret, Morgan, Toshiye, Colette, Martin, George, and Elinora—who give my life depth and meaning that extends far beyond the work I do. My father, whose discipline and tenacity is the part of me that generates the power to climb big mountains. My husband Dexter, who, over the course of finishing this PhD in a pandemic, was also my co-worker, my sounding-board, my dishwasher, my roommate, my confidant, my barber, my cheerleader, my alarm clock, my personal chef, my ride, my grocery delivery service, and my partner in all things. He is the invisible foundation buttressing everything I do. Finally, my best friend for nearly half my life, Erika: if you notice any heart in my work, it is only because she never stops pushing me to put it there. It is for your boys that I work toward a radically reimagined future.

### **CONTRIBUTION OF AUTHORS**

Chapter 1 was conceived and written by L. Delaney, and edited by J. Brown.

**Chapter 2** is an accepted publication (by the *International Journal of Plant Sciences* on 09/22/2021) conceived by B. Igić, who provided the preliminary data. With intellectual guidance from B. Igić, L. Delaney compiled the full dataset, performed the analyses, created the figures, and wrote the early drafts of the manuscript. B. Igić and L. Delaney wrote the final manuscript.

Chapter 3 was conceived and written by L. Delaney, and edited by J. Brown.

**Chapter 4** was conceived by B. Igić, who provided the preliminary data. With intellectual guidance from B. Igić, L. Delaney compiled the full dataset, performed the analyses, created the figures, and wrote the manuscript.

**Chapter 5** was conceived and written by L. Delaney, with intellectual guidance from J. Brown. J. Brown edited the manuscript.

**Chapter 6** was conceived and written by L. Delaney, with intellectual guidance from J. Brown. J. Brown edited the manuscript.

## TABLE OF CONTENTS

### **CHAPTER**

1	EVERYTHIN	IG EXISTING IN THE UNIVERSE IS THE FRUIT OF CHANCE AND
	NECESSITY	<sup>*</sup>
	1.1	Why is there something rather than nothing?
	1.1.1	Quantum state of affairs
	1.1.2	The matter with antimatter
	1.1.3	Elementary, dear data
	1.2	The fruit of chance6
	1.2.1	To make an apple pie from scratch
	1.2.2	Cherry bombs     7
	1.2.3	Through the grapevine
	1.3	The fruit of necessity
	1.3.1	Peas in a pod 10
	1.3.2	Inside the peach there's a stone
	1.3.3	Time flies like an arrow; fruit flies like a banana13
	1.4	The fruit of this work15
	1.5	<i>Omnia mea mecum porto</i>
	1.6	Cited Literature
2	THE PHYLC	GENETIC DISTRIBUTION AND FREQUENCY OF SELF-INCOMPATIBILITY
	IN FABACE	AE
	2.1	Abstract     26
	2.2	Introduction
	2.3	Materials and methods31
	2.3.1	Study system     31
	2.3.2	Breeding system data
	2.3.3	Chromosome counts and ploidy data
	2.3.4	Other trait data
	2.3.5	Phylogenetic relationships35
	2.3.6	Trait correlation analyses36
	2.3.7	Taxonomic name resolution37
	2.4	Results     38
	2.4.1	Breeding system in Fabaceae
	2.4.2	Trait associations43
	2.5	Discussion
	2.5.1	Features of SI in Fabaceae49
	2.5.2	Weak evidence regarding mechanism(s) of SI Fabaceae51
	2.5.3	Trait associations54

# TABLE OF CONTENTS (Continued)

2.6   Acknowledgements   58     2.7   Cited Literature   58     3   THE ORCHIDS AND THEIR BREEDING SYSTEMS   72     3.1   Cited Literature   87     4   SCIENTIFIC ADDENDUM TO THE ORCHIDS AND THEIR BREEDING SYSTEMS   92     4.1   Abstract   92     4.2   Materials and methods   93     4.2.1   Study system   93     4.2.2   Breeding system data   94     4.2.3   Other trait data   94     4.2.4   Phylogenetic relationships   96     4.3.7   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS q EXPLANATIONS OF ADAPTATION.   1     L AFRAMEWORK FOR SYSTEMATIC ANALYSIS   109     5.1   Abstract   109     5.2   Introduction   110     5.3.1   Conceptual framework: Epistemological and pedagogical aspects	<u>CHAPTER</u>			<u>PAGE</u>
2.7Cited Literature583THE ORCHIDS AND THEIR BREEDING SYSTEMS723.1Cited Literature874SCIENTIFIC ADDENDUM TO THE ORCHIDS AND THEIR BREEDING SYSTEMS924.1Abstract924.2Materials and methods934.2.1Study system934.2.2Breeding system data944.2.3Other trait data944.2.4Phylogenetic relationships964.2.5Taxonomic name resolution964.3Results974.3.1Index of self-incompatibility984.3.2Phylogenetic distribution of and frequency of SI984.3.3Other breeding systems1024.4Cited Literature1055UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION.1. A FRAMEWORK FOR SYSTEMATIC ANALYSIS1095.1Abstract1095.2Introduction1105.3.1Conceptual framework: Measurement1185.3.2.0Conceptual framework: Measurement1185.3.2.1The concept of adaptation1215.3.2.2Biological explanation in biology education1315.4Materials1345.4.1Participant sample1345.4.1Participant sample1345.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological explanations1395.5.3RQ3: Teleological versus teleonomic justification		2.6	Acknowledgements	58
3   THE ORCHIDS AND THEIR BREEDING SYSTEMS   72     3.1   Cited Literature   87     4   SCIENTIFIC ADDENDUM TO THE ORCHIDS AND THEIR BREEDING SYSTEMS   92     4.1   Abstract   92     4.2   Materials and methods   93     4.2.1   Study system   93     4.2.2   Breeding system data   94     4.2.3   Other trait data   95     4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of S1   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION   1     1. A FRAMEWORK FOR SYSTEMATIC ANALYSIS   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   The concept of adaptation   124 <		2.7	Cited Literature	58
3.1   Cited Literature   87     4   SCIENTIFIC ADDENDUM TO THE ORCHIDS AND THEIR BREEDING SYSTEMS   92     4.1   Abstract   92     4.2   Materials and methods   93     4.2.1   Study system   93     4.2.2   Breeding system data   94     4.2.3   Other trait data   95     4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.   1     1. A FRAMEWORK FOR SYSTEMATIC ANALYSIS   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   Theoretical background   118     5.3.1   Conceptual framework: Epistemological and pedagogical aspects <td>3</td> <td>THE ORC</td> <td>HIDS AND THEIR BREEDING SYSTEMS</td> <td>72</td>	3	THE ORC	HIDS AND THEIR BREEDING SYSTEMS	72
4     SCIENTIFIC ADDENDUM TO THE ORCHIDS AND THEIR BREEDING SYSTEMS     92       4.1     Abstract     92       4.2     Materials and methods     93       4.2.1     Study system     93       4.2.2     Breeding system data     94       4.2.3     Other trait data     94       4.2.3     Other trait data     95       4.2.4     Phylogenetic relationships     96       4.2.5     Taxonomic name resolution     96       4.3     Results     97       4.3.1     Index of self-incompatibility     98       4.3.2     Phylogenetic distribution of and frequency of SI     98       4.3.3     Other breeding systems     102       4.4     Cited Literature     105       5     UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.     1       1     Abstract     109       5.1     Abstract     109       5.2     Introduction     110       5.2.1     Research questions     117       5.3     Theoretical background     118 <t< td=""><td></td><td>3.1</td><td>Cited Literature</td><td>87</td></t<>		3.1	Cited Literature	87
4.1   Abstract.   92     4.2   Materials and methods   93     4.2.1   Study system   93     4.2.2   Breeding system data   94     4.2.3   Other trait data   95     4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION.   1     1   Affanterature   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   Theoretical background   118     5.3.2   Conceptual framework: Epistemological and pedagogical aspects   121     5.3.2.1   The concept of adaptation   124	4	SCIENTIF	IC ADDENDUM TO THE ORCHIDS AND THEIR BREEDING SYSTEMS	92
4.2   Materials and methods   93     4.2.1   Study system   93     4.2.2   Breeding system data   94     4.2.3   Other trial data   95     4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION.     I. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   Theoretical background   118     5.3.1   Conceptual framework: Measurement   118     5.3.2   Gonceptual framework: Measurement   121     5.3.2.1   The concept of adaptation   124		4.1	Abstract	92
4.2.1   Study system   93     4.2.2   Breeding system data   94     4.2.3   Other trait data   95     4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.   10     1. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   Theoretical background   118     5.3.2.1   Conceptual framework: Measurement   118     5.3.2.2   Biological explanation   121     5.3.2.3.2   Biological explanation   121     5.3.2.4   Conceptual understanding of natural selection   131     5.4   Materials		4.2	Materials and methods	93
4.2.2   Breeding system data   94     4.2.3   Other trait data   95     4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION.   1     1. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   Theoretical background   118     5.3.1   Conceptual framework: Measurement   118     5.3.2   Conceptual framework: Measurement   121     5.3.2.3   Biological explanation   124     5.3.2.4   Conceptual nderstanding of natural selection   131     5.4   Materials   134     5.4.1   Participant sample   134     5.4.2   Framework		4.2.1	Study system	93
4.2.3   Other trait data   95     4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION.   I. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.   109     5.1   Abstract   109   109   117     5.3   Theoretical background   118   117     5.3   Theoretical background   118   118     5.3.1   Conceptual framework: Measurement   118   118     5.3.2   Conceptual framework: Epistemological and pedagogical aspects   121     5.3.2.1   The concept of adaptation   124     5.3.2.2   Biological explanation in biology education   128     5.3.2.4   Conceptual understanding of natural selection   131		4.2.2	Breeding system data	94
4.2.4   Phylogenetic relationships   96     4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3.1   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.   1     I. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   Theoretical background   118     5.3.2.1   Conceptual framework: Measurement   118     5.3.2.2   Biological explanation   121     5.3.2.3   Biological explanation   124     5.3.2.4   Conceptual understanding of natural selection   131     5.4   Materials   134     5.4.1   Participant sample   134     5.5.1   RQ1: Validity and reliability   138     5.5.2   RQ2: Level of des		4.2.3	Other trait data	95
4.2.5   Taxonomic name resolution   96     4.3   Results   97     4.3   Index of self-incompatibility   98     4.3.2   Phylogenetic distribution of and frequency of SI   98     4.3.3   Other breeding systems   102     4.4   Cited Literature   105     5   UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.   109     5.1   Abstract   109     5.2   Introduction   110     5.2.1   Research questions   117     5.3   Theoretical background   118     5.3.1   Conceptual framework: Measurement   118     5.3.2   Conceptual framework: Measurement   121     5.3.2.1   The concept of adaptation   121     5.3.2.2   Biological explanation   124     5.3.2.2   Biological explanation   124     5.3.2.2   Biological explanation   124     5.3.2.2   Groeptual understanding of natural selection   131     5.4   Materials   134     5.4.1   Participant sample   134     5.5.5   Methods<		4.2.4	Phylogenetic relationships	96
4.3Results974.3.1Index of self-incompatibility984.3.2Phylogenetic distribution of and frequency of SI984.3.3Other breeding systems1024.4Cited Literature1055UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION.11. A FRAMEWORK FOR SYSTEMATIC ANALYSIS1095.1Abstract1095.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.2Conceptual framework: Measurement1185.3.2Silological explanation1215.3.2.3Biological explanation1245.3.2.4Conceptual understanding of natural selection1315.4Materials1345.5Methods1385.5.1RQ1: Validity and reliability1385.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		4.2.5	Taxonomic name resolution	96
4.3.1Index of self-incompatibility984.3.2Phylogenetic distribution of and frequency of SI984.3.3Other breeding systems1024.4Cited Literature1055UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.1091. A FRAMEWORK FOR SYSTEMATIC ANALYSIS1095.1Abstract1095.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		4.3	Results	97
4.3.2Phylogenetic distribution of and frequency of SI984.3.3Other breeding systems1024.4Cited Literature1055UNIVERSITY STUDENTS' DESCRIPTIONS § EXPLANATIONS OF ADAPTATION.1I. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.1095.2Introduction1105.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2.2Biological explanation1215.3.2.3Biological explanation1245.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications141		4.3.1	Index of self-incompatibility	98
4.3.3Other breeding systems1024.4Cited Literature1055UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.1095.1Abstract1095.1Abstract1095.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.2Conceptual framework: Measurement1185.3.2.1The concept of adaptation1215.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.5.4RQ4: Effects of previous exposure1385.5.4RQ4: Effects of previous exposure142		4.3.2	Phylogenetic distribution of and frequency of SI	98
4.4Cited Literature1055UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION. I. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.1095.1Abstract.1095.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2.2Biological explanation1215.3.2.3Biological explanation1245.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		4.3.3	Other breeding systems	102
5UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.I. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.1095.1Abstract1095.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.1The concept of adaptation1245.3.2.3Biological explanation1245.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.4RQ4: Effects of previous exposure142		4.4	Cited Literature	105
5.1Abstract1095.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.1The concept of adaptation1245.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1315.4Materials1345.4.1Participant sample1345.4.2Framework development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.4RQ4: Effects of previous exposure142	5	UNIVERS I. A FRAM	ITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION. IEWORK FOR SYSTEMATIC ANALYSIS.	109
5.2Introduction1105.2.1Research questions1175.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.1The concept of adaptation1245.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.1	Abstract	109
5.2.1Research questions1175.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.1The concept of adaptation1215.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.4RQ4: Effects of previous exposure142		5.2	Introduction	110
5.3Theoretical background1185.3.1Conceptual framework: Measurement1185.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.1The concept of adaptation1215.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.2.1	Research questions	117
5.3.1Conceptual framework: Measurement1185.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.1The concept of adaptation1215.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.4RQ4: Effects of previous exposure142		5.3	Theoretical background	118
5.3.2Conceptual framework: Epistemological and pedagogical aspects1215.3.2.1The concept of adaptation1215.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.4RQ4: Effects of previous exposure142		5.3.1	Conceptual framework: Measurement	118
5.3.2.1The concept of adaptation1215.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.3.2	Conceptual framework: Epistemological and pedagogical aspects	121
5.3.2.2Biological explanation1245.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.3.2.1	The concept of adaptation	121
5.3.2.3Biological explanation in biology education1285.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.4RQ4: Effects of previous exposure141		5.3.2.2	Biological explanation	124
5.3.2.4Conceptual understanding of natural selection1315.4Materials1345.4Participant sample1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.3.2.3	Biological explanation in biology education	128
5.4Materials1345.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.3.2.4	Conceptual understanding of natural selection	131
5.4.1Participant sample1345.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.4	Materials	134
5.4.2Framework development1345.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.4.1	Participant sample	134
5.4.3Questionnaire development1365.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.4.2	Framework development	134
5.4.4Administration1375.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.4.3	Questionnaire development	136
5.5Methods1385.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.4.4	Administration	137
5.5.1RQ1: Validity and reliability1385.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.5	Methods	138
5.5.2RQ2: Level of descriptions and explanations1395.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.5.1	RQ1: Validity and reliability	138
5.5.3RQ3: Teleological versus teleonomic justifications1415.5.4RQ4: Effects of previous exposure142		5.5.2	RQ2: Level of descriptions and explanations	139
5.5.4 RQ4: Effects of previous exposure		5.5.3	RQ3: Teleological versus teleonomic justifications	141
		5.5.4	RQ4: Effects of previous exposure	142

## TABLE OF CONTENTS (Continued)

### **CHAPTER**

5.6	Results
5.6.1	RQ1: Validity and reliability
5.6.2	RQ2: Level of descriptions and explanations
5.6.3	RQ3: Teleological versus teleonomic justifications
5.6.4	RQ4: Effects of previous exposure
5.7	Discussion
5.7.1	Limitations and future work
5.8	Cited Literature
UNIVER	SITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION.
II. A FRA	MEWORK FOR PEDAGOGICAL EXPLANATION.
6.1	Abstract
6.2	Background
6.2.1	Pedagogical explanation
6.2.2	Descriptions and explanations in teaching
6.2.3	Teaching adaptation
6.2.4	Model of belief change
6.3	Educational recommendations
6.3.1	LG1: Distinguish between descriptions and explanations
6.3.2	LG2: Identify processes that generate variation
6.3.3	LG3: Explain population-level features
6.3.4	LG4: Distinguish between teleonomic and teleological reasoning
6.4	Final thoughts
6.5	Cited Literature
ΑΡΡΕΝΓ	NCES
Anne	ndix A
Anne	ndix B
Δnne	ndix C
Appe	ndix C

## LIST OF TABLES

<b>TABLE</b>			<u>PAGE</u>
	Ι	SI frequency and breeding system coverage by Fabaceae subfamily.	39
	Π	Summary of results evaluating trait correlations in Fabaceae.	45
	III	Summary of SI distribution by orchid subfamily.	97
	IV	Site of fertilization failure following selfing in studied orchid species.	104
	V	Questions and approaches on floral color polymorphism.	126
	VI	Description of each course and student number by semester	135
	VII	Breakdown of grade level by major and semester.	135
	VIII	Categories for response coding with descriptions and examples.	140
	IX	Common teleological reasoning patterns.	142
	Х	Ultimate pattern descriptions and teleological responses by question.	150
	XI	Studies examining genetic phenomena possibly related to SI in Fabaceae.	212
	XII	Studies examining site-of-fertilization failure in Fabaceae species.	213

## LIST OF FIGURES

<b>FIGURE</b>		PAGE
1	Genus-level legume phylogeny displaying breeding system data coverage.	40
2	Distribution of ISI values in Fabaceae.	42
3	Site of fertilization failure following selfing in studied Fabaceae species.	44
4	Mosaic plot for best performing legume trait correlation model.	47
5	The orchid flower.	74
6	Botanical illustration of the yellow lady's slipper.	76
7	Botanical illustration of dioecous <i>Catasetum discolor</i> .	78
8	A cartoon depiction of purging.	84
9	Botanical illustration of autogamous <i>Ophyrus apifera</i> .	85
10	Genus-level orchid phylogeny displaying breeding system data coverage.	99
11	Proportion of each orchid genus that is SI or SC.	100
12	Distribution of ISI values in Orchidaceae from collected literature	101
13	Species-level orchid phylogeny with breeding system data.	103
14	A framework for organizing biological descriptions and explanations.	112
15	Using the framework to examine the adaptation of woody tissue.	116
16	Distribution of previously earned credits by student participant.	136
17	Overall distribution of student responses in both semesters.	146
18	Overall distribution of student responses by question in both semesters.	147
19	Relationship between levels and response type.	148
20	Reasoning type by question.	149
21	Reasoning type by descriptions versus explanations.	151
22	Proportion of legume species by genus that are SI.	214
23	Species-level Fabaceae phylogeny displaying SI and SC species.	215

## LIST OF ABBREVIATIONS

AIC	Aikake's information criterion
ACORNS	Assessment of Contextual Reasoning about Natural
	Selection
BSC	Biological species concept
CANS	Conceptual Assessment of Natural Selection
GSI	Gametophytic self-incompatibility
ISI	Index of self-incompatibility
LG	Learning goal
RI	Reproductive isolation
RQ	Research question
RSI	R-Nase-based self-incompatibility
SC	Self-compatibility
SI	Self-incompatibility
SSE	State-dependent speciation and extinction
UIC	University of Illinois at Chicago

#### **SUMMARY**

For thousands of years, humans have sought to understand the origins of our natural world and the vast complexity of life it contains. Some of the most compelling and powerful of the explanations put forth are not yet two centuries old. Still today we are documenting observed patterns in the distribution of traits and species with the goal of uncovering the processes that gave rise to them. This is the two-fold task of understanding the historical action of evolution: documenting patterns and investigating the causal processes that may generate such patterns.

This work contributes to that effort both directly and indirectly. First, as a direct contribution, I gathered data across two of the largest flowering plant families—Fabaceae and Orchidaceae, which together contain roughly 16% of all flowering plant diversity—to determine the frequency and phylogenetic distribution of one of the most powerful adaptations possessed by some of their species: the ability to genetically control their mating partners. While computational tools provide power to investigate large-scale models of evolutionary change, uncovering *true* historical processes is only possible with accurate and meaningful data. Thus I carefully evaluated more than a thousand published reports spanning three centuries, in nearly a dozen languages, ultimately creating databases for both families that each contain information for more than a thousand species—using this data to evaluate the current state of knowledge regarding the nature of this adaptation and its evolution in these groups.

This is how scientists make sense of natural selection. But despite the extraordinary advances made by evolutionary biologists since the publication of Darwin's masterpiece, natural selection remains one of the most poorly understood mechanisms of evolutionary change among the general public. As an

### SUMMARY (Continued)

indirect contribution, I next sought to examine how our students, the future scientists and leaders of the world, understand the historical process of natural selection and patterns of resultant adaptations. I developed a framework and a series of open-ended questions that I administered to nearly 600 student participants across all levels of biological study. I show how this framework has utility to both systematically assess student thinking and aid in developing instructional methodologies that address serious gaps in understanding.

Evolutionary theory is the explanatory framework that allows us to make sense of the natural world and the vast complexity of life it contains. Thus our understanding of evolution—as scientists, educators, students, members of our global community—is essential to our understanding and appreciation of the natural world.

### CHAPTER 1

### EVERYTHING EXISTING IN THE UNIVERSE IS THE FRUIT OF CHANCE AND NECESSITY

When in the height heaven was not named; And the earth beneath did not yet bear a name; And the primeval Apsū, who begat them; And chaos, Tiamat, the mother of them both. –The First Tablet, 1-4 (King, 1902)

Verily, in the beginning there was here the non-existent. -Sixth Kânda 1:1:1 (*Satapatha Brahmana*, 1897)

First there was the great cosmic egg. Inside the egg was Chaos, and floating in Chaos was P'an Ku, the Undeveloped, the divine Embryo. –Leach (1956), pg. 224

Tell it from the beginning, about what was generated first from among them all. First it was Chaos, and next broad-bosomed Earth. –Hesiod (2020), Lines 115-120

In the beginning, when God created the heavens and the earth, the earth was a formless void and darkness covered the face of the deep. -Genesis 1:1-2 ("The Book of Genesis," 1989)

The cosmos is all that ever is or ever was or ever will be. -Carl Sagan (1980), pg. 4

We are all of us here on Earth. In a universe of infinite wonder, we only know of a single planet that sustains life. And on this one planet, full of life, every creature can be traced back to one group of singlecelled organisms. Our deepest homologies—the 20 amino acids and four nucleotides we all share signify a profound connection between us that transcends the artificial boundaries that humans create. Which is to say that all connection is a matter of scale. But we humans are compelled—have always been compelled—to make sense of this connection, to understand why we are here and divine meaning from our existence: to understand the provenance of Earth and the exceptional life it cradles. This thesis is about one small piece of that pursuit, about how we understand natural selection; its introduction is about *why we try*. It is carving out a small space for a personal statement on the nature of things. It is about the big questions that we reach toward incrementally, without knowing if we will ever reach the answers. It is a collection of musings on why the pursuit of understanding natural selection has everything and nothing to do with natural selection itself. Everything, as every bit of us has something to do with natural selection, as we are sculpted by it like every creature that has lived on Earth. Nothing, as our impulse to know the universe and our place in it, to tell our *true* origin story, involves exposing a long and tangled causal history that only somewhat recently involves natural selection at all. And of course, natural selection alone cannot fully account for our origins. So first we begin by placing natural selection in context. First we begin at the beginning of all things.

#### 1.1 Why is there something rather than nothing?

Why is there something rather than nothing? can of course mean a great deal of things (Brenner, 2016; Leibniz, 1989). Heidegger (2000) famously calls this the first, the broadest, the deepest, the most original and central question of all philosophy. But here I mean it in the most literal and unanswerable sense: why, in the beginning, around 13.8 billion years ago, was there was there suddenly something? This tiny egg we now call a singularity, all matter and energy condensed into a single point, all the potential for big stars and little squirrels, that gave birth to the universe? And if all matter and energy were borne of the singularity, what did it exist inside? What happened to precipitate its outward expansion? This "big bang," that in the void made no sound, that began smaller than an atom. The multiverse theory suggests that perhaps our universe is part of a larger collection of universes—that our

universe is simply a ramet that sprouted from a larger genet (Linde, 2017). But then, what of the first universe? Into *what* exactly is the universe expanding?

Origin stories cannot tell us what exists outside the bounds of now, what came before. Our knowable universe did not exist until the Big Bang, so who knows what was before, or how long it was there. An old saying goes: "There's speculation, then there's more speculation, then there's cosmology" (Kaku, 2005). Perhaps as far as we are concerned there really *was* nothing. Not big empty space, not vast amounts of darkness, not suspended primordial particles, biding their time: just nothing. The absence of all things. The only thing we can say for sure is that from this unknowable nothingness sprang something.

Today we mostly take for granted what exists, from photons to phytoplankton. We rarely stop to think about why there should be any universe at all.

#### 1.1.1 Quantum state of affairs

It took a trillionth of a second of knowable time for the four fundamental forces that shape our physical world to develop, these physical laws that govern and constrain all matter and energy in our universe. Why should these forces have developed as they did? If our singularity were slightly altered in composition, if the first moments of the universe were less hot or more hot by a fraction of a degree, if gravity were slightly weaker, if one infinitesimally small interaction were altered in that first trillionth of a second, would that change the subatomic particles that make up everything we know? Perhaps a universe with radically unfamiliar physics would exist now. Or perhaps there really is only one stable possibility. Perhaps any other kind of universe would simply collapse on itself and begin again.

But as it happens the universe does not collapse. In this first trillionth of a second, like so many of our earliest origin stories describe, Chaos: the universe expands rapidly an octillionth  $(10^{27})$  in every direction in an event known as inflation (Martin, 2019). The speed and magnitude of this expansion boggles the mind, like an electron becoming the size of the solar system faster than the blink of an eye. Such chaotic moments in the early universe violate all currently-understood physical laws. But it is now thought that this impossible expansion and cooling is what truly precipitated our knowable universe: from the resulting quark-gluon plasma, the final forces of constraint begin to take shape.

#### 1.1.2 The matter with antimatter

It is not clear why there is any matter at all. In the hot plasma of the inchoate universe, both matter and antimatter—similar particles with opposite charges—are created copiously in pairs. As the plasma cools with the expansion of the universe, the production of pairs slows, and matter and antimatter should ultimately annihilate with each other (Boucenna & Morisi, 2013). This is what such particles do when they come into contact: the only thing left behind in their wake is pure energy. But this is not a universe of pure energy. For all the "stuff" we see in the universe, from every nebula to every comet, from every planet in our solar system to every human-made artifact on Earth: *something* is there. Something is there because for every billion pairs of matter and antimatter created in the early universe, two matter particles were not annihilated. Two matter particles remained.

This great upset of equilibrium should not have occurred. If matter and antimatter are created in pairs that annihilate on contact, then something must account for the observed asymmetry: physical laws again must be violated (Canetti et al., 2012). Thus the remnants of this cosmic imbalance—this

relatively small and lucky collection of atoms that we call "something"—should not really be here. And yet.

#### 1.1.3 Elementary, dear data

The data is clear: scientists know the first and oldest elements created after the Big Bang. Hydrogen, helium, and a smattering of lithium are formed as the universe continues to cool and electrons can finally bind with protons. But for as much as the next several hundred million years, we know of nothing more than the cosmic background radiation leftover from the processes of recombination and decoupling that made these very first elements (Bond & Efstathiou, 1987). We do not know how the first galaxies and stars formed, only that such large structures first require dense regions of gas. It is puzzling that in a mostly homogeneous early universe such dense pockets of gas would form at all: cosmologists theorize that tiny quantum fluctuations during inflation are responsible for the uneven distribution of matter needed to form larger galactic structures (Dayal & Ferrara, 2018). These minuscule and random fluctuations serve as the architect for the grand structure of the cosmos as we observe it today.

Somewhere between several hundred million and a billion years after inflation the first galaxies and stars appear in the universe. Stars, these instruments of elemental conception. Is their work of fusion simply inevitable with dense pockets of hydrogen gas and hundreds of millions of years? Perhaps not. All we know is that without stars, there would be no elements heavier than beryllium in the universe. No aluminum for our foil or neon for our signs, no sodium or potassium for ion channels in our cells, no carbon to build any life at all. Such elements are all forged through the life and death of stars. The earliest among them—likely monumental beasts, more than a hundred to a thousand times the mass of our sun—fuse hydrogen nuclei to create helium before exploding in massive supernovae as they their

exhaust hydrogen supply. These explosions enrich the surrounding area with metallic elements for the next four million years (Heger & Woosley, 2002). Now new stars may be born, different stars. As infants the stars are made of mostly hydrogen or helium, but eventually some fuse helium to make beryllium, fuse beryllium to make oxygen—those that explode create heavier metallic elements that make up the familiar rocks we observe in space. Including, naturally, the most important rock of all.

#### **1.2** The fruit of chance

The light that reaches us from distant galaxies has traveled for millions, sometimes billions, of years. Thus, as we peer into the universe we also look back into time. Any civilizations that may have existed within such galaxies—perhaps even the galaxies themselves—have likely returned to the interstellar dust from which they came. Given such big distances, such long timescales, we cannot say exactly how miraculous it is that life appeared in our solar system. We cannot say exactly how it happened at all: only that a great deal of it happened by chance.

#### **1.2.1** To make an apple pie from scratch

The Milky Way is old. Very old. Its estimated age makes it among the first galaxies to form 13 billion years ago (Pasquini et al., 2004). Like many galaxies, it exists as it does because of the stochastic quantum fluctuations that were magnified when the early universe inflated. But not all galaxies are capable of supporting complex biological life. The formation of a terrestrial planet of a suitable size in a suitable location—away from gamma ray bursts, globular clusters, frequent supernovae, and Oort cloud comets, among other things—appears to depend primarily on the location and composition of the sun around which the planet forms (Gonzalez et al., 2001). In other words, the pocket of interstellar

dust that birthed our solar system determined—long before our sun even existed—whether or not a terrestrial planet could form at all.

We locate the oldest stars in the universe by finding those that are are composed entirely of the only elements that existed then: hydrogen and helium, with trace amounts of lithium. Such stars— population III stars—contain none of the metallic elements that come from older stars' supernovic leftovers. Our sun is a population I star, the youngest group, meaning that it was borne of interstellar dust richer in heavier elements, created by the death of far more ancient stars. This is crucial because the stuff that made our sun is also what makes its planetary bodies: there could be no terrestrial planets without first these heavy elements (Fischer & Valenti, 2005). Strangely, our solar system's composition of inner planets is unlike most others that we observe: often the inner solar system is tightly packed with large uninhabitable and non-terrestrial "superearths" up to fifty times the size of our home planet (Batygin & Laughlin, 2015; Schulze-Makuch, 2017). Which is to say that it seems our relatively uncluttered inner solar system is rare.

#### 1.2.2 Cherry bombs

Several billion years after the Milky Way forms, a large pocket of hydrogen gas, more than 25,000 light-years from galactic center, begins to collapse. The outward pressure exerted by the gas is overwhelmed by the inward pressure of gravity and a protostar is born. Around this protostar a disk of particles starts to spin (Greaves, 2005). A period of coalescence takes place with larger and larger rocks smashing together with wicked force. Of course, from all the smashing comes planets. The planets of our inner solar system were long thought to emerge first during this period, followed later by the planets of the outer solar system. But then computers got involved, taking us beyond the boundaries of our own big brains. Cosmologists discovered that in every computational simulation of early planet formation there was one consistent problem: not a single one could recover Mars. The models repeatedly predicted that a planet 5- to 10-times the size of Mars should exist in its location. So pervasive this conundrum that it became referred to simply as *the small Mars problem* (Raymond et al., 2009).

The best explanation for the size of Mars (and several other features of our unusual inner solar system) involves a violent dance between Jupiter and Saturn known as the *Grand Tack* (Raymond & Morbidelli, 2014). The story goes like this: at one time in the inner solar system, there exist several superearth-like planetary bodies orbiting close to our sun. But Jupiter, accumulating in mass, migrates toward the sun, disrupting the orbit of every planet in its wake and causing a collisional cascade of obliteeration. Much of the mass of these early superearths hurtles into the sun. Meanwhile, Saturn catches up to Jupiter, trapping it in orbital resonance: both planets then again change tacks and migrate outward together. What remains? Venus and Earth will form from the remnant rocks of the lost superearths that were pushed inward by Jupiter. The region now occupied by Mercury and Mars contains little of this debris and forms our smallest terrestrial planets. The stage is set for an unusually peaceful place in the galaxy.

But first. Earth is not stable: fluctuations in its rotation would eventually cause severe seasonal extremes problematic for imminent life (Cockell et al., 2016). Around when the rocks that comprise the Earth were mostly accumulated and molten, as they traveled in orbit around the sun, a collision occurred by chance with a planetary body that was roughly the size of Mars (Cameron & Ward, 1976). The remnants of this collision stabilize the Earth's rotation and can still be seen today with the naked eye. We call it the moon.

#### 1.2.3 Through the grapevine

We know the chemical building blocks of life. We have diverse hypotheses about how the structure and composition of these molecules suggests the potential for biological organization. Nucleotides form long chains of distinct subunits with complementarity. Lipids form micelles in water. We know that the existence of life—this special kind of "something"—likely occurred not all at once but in many stages of increasing complexity. Life, from Chaos. A Divine Embryo. But we cannot say exactly how it occurred. Nor can we say how the necessary chemical building blocks appeared on our planet. Nor can we say if another style of biochemistry, one not built on carbon and water, is possible. All we know is that at some point around four billion years ago, evidence of microorganisms appears in underwater hydrothermal vents (Dodd et al., 2017). This extraordinary life—life that easily might never have existed, without just the right quantum fluctuation at just the right moment to seed the Milky Way, that ultimately contained just the right interstellar medium in just the right place for our perfect sun, around which our magically protected rock flies, that rock which by all accounts should not really be there.

#### **1.3** The fruit of necessity

The planets exist where they do because they must: it could be said that the sun invokes enormous selective pressure on its surrounding matter. There is a reason why gas giants develop at a distance past the sun astronomers call *the frost line* (Fisher et al., 2018). There is a reason why the only planet in close orbit to our star is mostly made of iron. Which is not to say that every planet next to every star is like Mercury, only that what's there is the only stuff that could be there, given the environment. This is one kind of fruit of necessity. But planets and their matter do not have the constraint of heritability. When

stars go supernova, the new star that forms is not bound irrevocably to the old. The star that may form in its dust can be a very different kind of star because the dust itself is different.

In nature, nothing is born anew. We are not created from scratch. Every creature is constrained not only by the surrounding environment—like all matter in the universe—but also by its ancestors. This is when the fruit of necessity becomes truly compelling. Now the fun stuff begins.

#### 1.3.1 Peas in a pod

On early Earth, there is no desert nor no tundra. There is barely any oxygen in the air at all. As macromolecules spontaneously organize into structures, it is not clear when selective pressure becomes biological, or in other words, natural. Nor is it clear exactly which macromolecules organized first and into what flavor of structure. Passionate debates among scientists cover a wide landscape and are often diametrically opposed (Luisi, 2016). But in truth there are many possibilities as to how non-living entities organized into living ones. We do know that RNA possesses special and peculiar features that implicate it in the origin of life, as the first organized and self-replicating structure. Because of RNA's unique ability to both store information and catalyze reactions, the "prebiotic RNA world" has been called "the molecular biologist's dream" (Joyce & Orgel, 1993). But the dream is not shared by everyone: it has also been called "the prebiotic chemist's nightmare" (Robertson & Joyce, 2012).

While the intellectual leap from self-replicating strands of RNA to the intricate ballet of DNA, mRNA, tRNA, and protein construction at first appears comparatively short, there remain many unsolved chemical problems that leave this path to life shrouded in shadows. We know that chemically speaking, an RNA world requires the existence of at least one RNA polymer long enough to trigger polymerase activity. But we also know that the statistical chance of forming such a macromolecule in the absence of enzymes is nearly zero (Totani, 2020). Moreover, however many activated nucleotides there are on early Earth, waiting to become polymers, such a supply would eventually be depleted without continued biosynthesis (Robertson & Joyce, 2012). The chicken and the egg is an ancient ditty long-solved. Of course, the amniotic egg came first by nearly 200 million years. Perhaps instead we should wonder: how do we get RNA macromolecules, if protein synthesis requires RNA and RNA synthesis requires proteins? Recent work demonstrates that minor alterations in tRNA—the molecule that links nucleotides to amino acids—impart an ability for autonomous replication (Kühnlein et al., 2021). Such an ability would suggest that tRNA forms the first bridge between the RNA and DNA worlds. Of course, it does not explain how such a molecule arose in the first place. Nor does it fully explain the emergence of functioning and living cells.

Some hypothesize that compartments originated before replicators (Monnard & Walde, 2015). In part, because compartmentalization is a major unifying feature of living organisms. In part because this feature facilitates the occurrence of chemical reactions in tandem without interference, and dynamic but highly ordered interaction with the surrounding environment. Thus compartments may provide means to carry out somewhat protected metabolic reactions. Moreover, the behavior of lipids in aqueous environments—that is, the ability to spontaneously aggregate and form micelles, vesicles, droplets, and bilayers—provides an obvious path to the formation of early "protocells" (Segré et al., 2001). At first such structures would likely be subject to the simple selective pressures that physical laws and surrounding environments place on matter: some collection of these early protocells "survive" while others do not, but the differences between them are not heritable. Without heritability the specific molecular components that allow a protocell to best capture energy from its environment cannot be maintained, because the instructions for such components cannot be transmitted. Components organize, capture energy, and disappear. But perhaps by chance, one develops a method of storing a small piece of information that can be used to replicate some particular component. Of course, any protocell that can use stored information, even poorly—most especially information that provides *any kind* of advantage—will increase greatly in frequency. It must, because it's the only game in town. Over long timescales with sustained interaction between protocells, these processes would be refined as more instructions are built: the information capacity of four complementary nucleotides is enormous (Eigen, 1971). But of course, this does not explain the formation of metabolism that facilitates energy capture. Nor does it fully explain the emergence of functioning and living cells.

If life emerges from an information-only prebiotic replicator world, the very nucleotides that represent our deepest homology may themselves be the product of cardinal selection (Subramanian et al., 2020). Instead if compartmentalization and metabolism occur prior to information transfer, natural selection does not act until molecular machinery becomes heritable. This matter is not trivial, because it informs an essential piece of the puzzle, of how and when natural selection truly begins influencing the formation and structure of living organisms.

#### **1.3.2** Inside the peach there's a stone

Are we inevitable? Is this moment inevitable? Just what proportion of the probability of you you reading this sentence right now was decided in that first trillionth of a second that the universe inflated? The views put forth on such questions run the ostensible gamut, from the irrefutable and deterministic certainty of human origins to the near impossibility of our existence. Stephen J. Gould writes, "And so ultimately, the question of questions," by which he means, the inevitability of human origins, "boils

down to the placement of the boundary between predictability under invariant law and the multifarious possibilities of historical contingency" (1990, Chapter IV). Gould argues for a relatively low boundary—famously, he argues that if the tape of life were rewound and replayed, human beings are unlikely to originate again—and I tend to agree. But how we formulate answers to such a question depends heavily on what we mean when we say "inevitable."

Like all things, the notion of inevitability itself depends, in part, on the nature of the universe and the mechanics of inflation. Do we exist inside a single fixed universe? Or rather, do we exist inside a universe that is merely one piece of a grander collection, an infinite array of universes? If the multiverse theory is true, if indeed there exist infinite universes that vary in their construction and physical laws, then two formerly dichotomous opinions can be simultaneously true: that our existence is both inevitable and exceedingly unlikely. Why? Because all histories may be repeated *ad infinitum* in an infinite array of multiverses. Given eternity, the emergence of complexity is simply unavoidable across the system as a whole (Koonin, 2007). But to be the product of complexity in *that one* universe—well, that is nothing short of miraculous.

#### 1.3.3 Time flies like an arrow; fruit flies like a banana

The magically protected rock we live on affords life eons to evolve. And eons can produce some truly spectacular results. To look at the fruit of natural selection today—the enormous variety of living organisms uniquely suited to their environments—and to see the connections between these organisms, it becomes difficult to appreciate that life on Earth was not always exactly as we have known it. Which is to say that we cannot always trust our own perception of *what is*. Our instincts, our intuitions, are shaped over time by their capacity to aid in our survival. Information processing is a trait

like any other: those that process information in a manner that increases their fitness will increase in the population. Meaning that over generations, human beings evolved extraordinary abilities to perceive and manipulate the surrounding environment, inasmuch as it benefits our continued success in the population.

Atoms are mostly empty space. What I look at when I look out my window is mostly nothing. So why do I see something? I see something because seeing mostly empty space is unlikely to make survival on this planet possible. Seeing something is the fruit of evolution. And even only on Earth, there are many ways of "seeing": just ask a bat (Nagel, 1974). Yet in one important sense, our perceptions reflect the universe accurately because such perceptions produce tangible results. We derive mathematical equations that *must* account for true physical laws, because those equations build very real bridges and skyscrapers, rocket people to the moon and back. We build instruments to extend and enhance our perceptions, gathering true data on the structure and composition of our sun, using that knowledge to forge weapons with palpable and ugly consequences. The data *must* be meaningful, because the data works. But the data also shows us where our human perceptions fail. The data demonstrates that there are a good many aspects of this universe that go against our instincts or that we cannot perceive at all. To get at what is Real or True, we must go far beyond accepting what we perceive as representative of *what is*.

The theory of evolution is counter to many constitutional human intuitions, and scratches at one of our most profound insecurities: uncertainty. How can there be meaning and purpose in a universe we arrived in by chance? Put another way, what is the purpose of meaning when the meaning is derived only through our big brains? Which is to say that the concept of meaning we cherish so deeply is one of our own making. But perhaps such an understanding may be a gift, albeit a different kind of gift than the belief we were created on purpose, for a purpose. Perhaps it demonstrates why we must be good to each other and the fragile life with which we share this special rock. Perhaps it brings us back where we belong—not exalted above any other microbe or flower that inhabits the Earth—but just another thread in the rich and dependent tapestry of life woven by chance and necessity. This life that might easily never have appeared. This life we know, exactly as it is because it had to be, given everything that came before. This life that today is largely in our care.

#### **1.4** The fruit of this work

If natural selection runs contrary to our human intuitions, why bother with the work it takes to override such intuitions? Does it matter that we understand *why* prokaryotic cells and eukaryotic cells are different, if we understand *that* they are? Perhaps it is tempting to think that the acquisition of descriptive knowledge is enough "education" for our students, but I suspect that there is much more at stake than whether or not graduates know that the mitochondria is the powerhouse of the cell. In truth, the educational systems of this country—by focusing largely on content at the expense of cultivating metacognitive and higher-order reasoning skills—are not designed to produce empowered critical thinkers. Moreover, access to quality education and educational opportunity is neither equal nor equitable, which is to say that such systems are not designed to support (certain) students at all (Alexander, 2012; Love, 2019). When we consider the exercise of meaningful education, hopefully most agree that it must move beyond the realm of rote fact memorization. It must aim to impart greater truth, greater wisdom—and the ability to continue the exercise of learning well beyond the classroom. Meaningful education is essential to civic participation. It is essential to the maintenance of a just and peaceful society. And it is essential to interrupting generational poverty (Bloome et al., 2018).

All of biological sciences is built upon the foundation of evolutionary mechanisms. Teaching biology without the foundation that undergirds the discipline moves biology from a profound science that necessitates robust critical thinking skills to a descriptive exercise of rote memorization. How many times have we heard our students lament, *Biology is just memorization*? More often than not, these students are right. Evolutionary theory is routinely omitted from large swaths of their biological education. Instead of wrestling with the complex interplay between the overlapping stochastic and deterministic processes that shape the natural world, students are regurgitating the steps of the citric acid cycle. This is not meaningful education. This will not help students make sense of complexity in other areas of study. Not only does it emphasize low-level cognitive skills, but it fails to demonstrate why biological facts are significant.

Natural selection tells us that heritable traits that increase an organism's ability to survive or reproduce will increase in frequency in a population over generations. We know *why* organisms possess adaptations, but it is often unclear *how* the mechanisms of evolution interact in practice over time. Natural selection is only possible with existing and heritable variation. It works not on optima but on realized genetic and phenotypic features. How much of species' evolutionary fates are determined by stochastic processes—those that underlie the production of variation, or those that result from offspring produced in finite populations? How do such stochastic processes interact with deterministic ones like selection? To understand such historical processes, we first need to gather data on the distribution of adaptations across groups. In the first half of this thesis, I examine an adaptation with profound effects on the evolution of flowering plants: self-incompatibility (SI), a suite of genetic mechanisms that prevent self-fertilization in otherwise fertile hermaphroditic plants. It is not known just how many SI systems exist or their evolutionary relationships. We do know that SI systems influence the amount and the distribution of genetic variation in space and over time by controlling the genetic relatedness of mating pairs. This has broad consequences for the action of natural selection, a force constrained by existing variation, and thus the distribution of all other traits. One particular flavor of SI—S-RNase-based GSI (RSI)—is found across several distantly-related families and believed to be ancestral to core eudicots (Igić et al., 2008). RSI is also the most well-characterized SI system. Problematically, there remains a surprising paucity of data concerning the phylogenetic distribution and frequency of SI across flowering plants. Data of this kind is essential in determining the selective pressures influencing the origin and maintenance of these systems across macroevolutionary timescales, and determining how these systems influence the diversity and ecological success of angiosperms as a group.

These first chapters are concerned with the collection of meaningful data on SI and its distribution across two of the largest angiosperm families: Fabaceae, nested within core eudicots and long hypothesized to possess RSI, and Orchidaceae, which together contain roughly 50,000 species. The ecological and economic significance of both families has motivated an enormous output of research concerning the reproductive biology of their species. Yet characterization of SI is no simple exercise, requiring careful and laborious pollination experiments—simply self-pollinating a plant and declaring it SI when it does not set seed is insufficient. What if the stigma (female reproductive organ) was not yet receptive? What if the pollen (male gametes) had not yet matured? What if generations of outcrossing had resulted in inbreeding depression from sheltered recessive mutations? Thus I examined over a thousand published reports—empirical papers, flora and field guides, monographs, conference proceedings, and theses—using this information to create publicly-available datasets on SI and related traits. **Chapter 2** summarizes the data gathered from Fabaceae and evaluates the hypothesis that RSI operates in this group. **Chapter 3** is directed at a nonscientific audience of orchid enthusiasts to introduce the concept of breeding systems and their consequences for flowering plant evolution. **Chapter 4** serves as an addendum to this chapter, summarizing my research findings on SI and its distribution in Orchidaceae.

We know why organisms appear uniquely designed for their environments—but how many of our students can explain the products of natural selection? Imagine the cognitive skills necessary to understand something like the evolutionary forces shaping the diversity of flowering plants. First, I must recognize the observed pattern of variation at the individual level—I must recognize, as so many failed to before Darwin, the incredible variation that exists within species. I must understand the proximate processes that account for such variation at the individual level, like mutation, sexual reproduction, and recombination. But the diversity of *all flowering plants* is an emergent feature that is not immediately obvious from these proximate-level patterns and processes. I must also recognize that what I see now is not always what was there. As Mayr (1961) famously notes, there is another layer of causality in biology: the ultimate or evolutionary. Thus I must recognize the ultimate species-level pattern that arises from the long-term process of organisms taking advantage of heterogeneous environments in different ways. There are biological tradeoffs to be made in this process, resulting in the production of a great variety of "ways to live."

Biology education, on the whole, is not preparing students with the higher-order cognitive skills required to understand evolution and make sense of the discipline of biological sciences. Far too often, siloed courses across subdisciplines present fact after fact without demonstrating for students how these facts connect and fit within the broader context. Are we teaching our students how to explain rather than describe facts? Again: how many of our students can explain the products of natural selection? Very few empirical studies examine whether or not students more commonly provide factual descriptions or causal explanations when asked causal biological questions. Chapter 5 presents the development of a questionnaire and framework used to systematically assess student thinking as it relates to the concept of adaptation. The framework allows us to easily categorize the kinds of responses students provide—descriptions or explanations at the individual or species level. In a study spanning four courses, from introductory to advanced, including nearly 600 student participants, I find that students overwhelmingly prefer species-level descriptions to explain adaptation—regardless of their prior preparation and performance in the course. This result indicates that a major issue in understanding adaptation is the ability to form causal explanations that integrate reasoning from different levels of biological hierarchy. Chapter 6 demonstrates how the framework developed in Chapter 5 may be used in the classroom to help students work on developing the important higher-order cognitive skills necessary to make sense of natural selection: namely, the ability to integrate facts appropriately and formulate causal explanations.

#### 1.5 Omnia mea mecum porto

In the early spring of 1987, a sperm found an egg. This is the start of my origin story. Any minute change in the chain of events from the Big Bang to this moment may have meant that a different sperm

found a different egg, or that no sperms or eggs existed at all. Richard Dawkins wrote, "However many ways there may be of being alive, it is certain that there are vastly more ways of being dead" (1996, Chapter 1). I rather think that it is far likelier to be nothing at all than something, alive or not. But events unfolded exactly as they did and *that* sperm found *that* egg, and I was born nine months later on the day before Christmas. I had not yet lived a full year when the final model of the Apple II series, the Apple IIc Plus, was released.

I type this today from my third-generation MacBook Air, with more than 6,000 times the computing power than the Apple IIc Plus. It also weighs a lot less. This spectacular and human device, that harnesses the flow of electrons in switches that represent the binary code underlying all of my graduate work and everything contained in this thesis. The transistors performing this work are made of silicon that—like me and everything I have ever known or held or continue to carry with me—was borne of a star somewhere in the universe. How remarkable. That this one moment of space and time occurred, and that not only did it occur but that it did so in exactly this way, this way that allows me to sit here at *this* computer typing *these* sentences, on a planet such as this. That, in a way, our sentience, the consciousness we bring to bear on the nature of our universe's existence, is the ultimate meaning we desperately seek, an incredible product of the long action of natural selection. The Big Bang the apotheosis of origin stories, our nascent universe the Goldilocks of matter and energy. Every "something" contained within it connected as a causal descendant through the same sequence of stochastic and deterministic events; all fruits of the same tree. And we the narrators of this improbable story, the curious detectives that give consequence to these wondrous events.

#### 1.6 Cited Literature

Note that the title and several subtitles were inspired by or come from other works: the main title from Monod (1971); 1.1.3 from Bowman (1988); 1.2.1 from Malone (1980); 1.2.3 from Whitfield & Strong (1966); 1.3.2 from Atwood (1997); 1.3.3 is widely attributed to Groucho Marx.

Alexander, M. (2012). *The New Jim Crow: Mass incarceration in the age of colorblindness*. The New Press. Atwood, M. (1997). *Alias Grace*. Anchor House.

- Batygin, K., & Laughlin, G. (2015). Jupiter's decisive role in the inner Solar System's early evolution. Proceedings of the National Academy of Sciences, 112(14), 4214–4217.
- Bloome, D., Dyer, S., & Zhou, X. (2018). Educational inequality, educational expansion, and intergenerational income persistence in the United States. *American Sociological Review*, *83*(6), 1215–1253.
- Bond, J., & Efstathiou, G. (1987). The statistics of cosmic background radiation fluctuations. *Monthly Notices of the Royal Astronomical Society*, *226*(3), 655–687.
- Boucenna, S. M., & Morisi, S. (2013). Theories relating baryon asymmetry and dark matter: A mini review. *arXiv Preprint arXiv:1310.1904*. http://arxiv.org/abs/1310.1904
- Bowman, R. (1988). Elementary, dear data. Star Trek: The Next Generation. USA: Paramount.
- Brenner, A. (2016). What do we mean when we ask "Why is there something rather than nothing?". *Erkenntnis*, *81*(6), 1305–1322.
- Cameron, A. G., & Ward, W. R. (1976). The origin of the moon. *Lunar and Planetary Science Conference*, 7.
- Canetti, L., Drewes, M., & Shaposhnikov, M. (2012). Matter and antimatter in the universe. *New Journal of Physics*, *14*(9), 095012.

- Cockell, C. S., Bush, T., Bryce, C., Direito, S., Fox-Powell, M., Harrison, J., Lammer, H., Landenmark, H., Martin-Torres, J., Nicholson, N., & others. (2016). Habitability: A review. *Astrobiology*, *16*(1), 89–117.
- Dawkins, R. (1996). The Blind Watchmaker: Why the evidence of evolution reveals a universe without design. WW Norton & Company.
- Dayal, P., & Ferrara, A. (2018). Early galaxy formation and its large-scale effects. *Physics Reports, 780*, 1–64.
- Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., O'Neil, J., & Little, C. T. (2017). Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature*, *543*(7643), 60–64.
- Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Die Naturwissenschaften*, *58*(10), 465–523.
- Fischer, D. A., & Valenti, J. (2005). The planet-metallicity correlation. *The Astrophysical Journal*, *622*(2), 1102.
- Fisher, C., Hart, P., Mallon, M., Lyall, A., & Irvine, I. (2018). Thermal infrared spectral comparisons of minerals and asteroids in relation to the frost line. *Lunar and Planetary Institute*. https://www.lpi .usra.edu/exploration/education/hsResearch/presentations/2017-2018/posters/Waynflete -Asteroid.pdf
- Gonzalez, G., Brownlee, D., & Ward, P. (2001). The galactic habitable zone: Galactic chemical evolution. *Icarus*, *152*(1), 185–200.

Gould, S. J. (1990). Wonderful Life: The Burgess Shale and the nature of history. WW Norton & Company.
- Greaves, J. S. (2005). Disks around stars and the growth of planetary systems. *Science*, *307*(5706), 68–71.
- Heger, A., & Woosley, S. E. (2002). The nucleosynthetic signature of population III. *The Astrophysical Journal*, *567*(1), 532.
- Heidegger, M. (2000). Introduction to metaphysics. Yale University Press.
- Hesiod. (2020). Cosmogeny. In G. Nagy (Trans.), *Theogony*. The Center for Hellenic Studies. https:// chs.harvard.edu/primary-source/hesiod-theogony-sb/
- Igić, B., Lande, R., & Kohn, J. R. (2008). Loss of self-incompatibility and its evolutionary consequences. *International Journal of Plant Sciences*, *169*(1), 93–104.
- Joyce, G. F., & Orgel, L. E. (1993). Prospects for understanding the origin of the RNA world. *Cold Spring Harbor Monograph Series*, *24*, 1–1.
- Kaku, M. (2005). Parallel worlds: A journey through creation, higher dimensions, and the future of the cosmos. Doubleday Books.
- King, L. W. (1902). The seven tablets of creation.
- Koonin, E. V. (2007). The cosmological model of eternal inflation and the transition from chance to biological evolution in the history of life. *Biology Direct*, *2*(1), 1–21.
- Kühnlein, A., Lanzmich, S. A., & Braun, D. (2021). tRNA sequences can assemble into a replicator. *Elife*, *10*, e63431.
- Leach, M. (1956). The beginning: Creation myths around the world.
- Leibniz, G. W. (1989). The principles of nature and of grace, based on reason. In *Philosophical papers and letters* (pp. 636–642). Springer.

Linde, A. (2017). A brief history of the multiverse. *Reports on Progress in Physics*, 80(2), 022001.

- Love, B. (2019). We Want to Do More Than Survive: Abolitionist teaching and the pursuit of educational *freedom*. Beacon Press.
- Luisi, P. L. (2016). *The emergence of life: From chemical origins to synthetic biology*. Cambridge University Press.
- Malone, A. (1980). The Lives of Stars. Cosmos. UK: British Broadcasting Corporation.
- Martin, J. (2019). Cosmic inflation: Trick or treat? Fine-Tuning in the Physical Universe.
- Mayr, E. (1961). Cause and effect in biology. *Science*, *134*(3489), 1501–1506.
- Monnard, P.-A., & Walde, P. (2015). Current ideas about prebiological compartmentalization. *Life*, *5*(2), 1239–1263.
- Monod, J. (1971). *Chance and Necessity: An essay on the natural philosophy of modern biology*. Alfred A. Knopf.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83(4), 435–450.
- Pasquini, L., Bonifacio, P., Randich, S., Galli, D., & Gratton, R. (2004). Beryllium in turnoff stars of NGC
  6397: Early Galaxy spallation, cosmochronology and cluster formation. *Astronomy & Astrophysics*, 426(2), 651–657.
- Raymond, S. N., & Morbidelli, A. (2014). The grand tack model: A critical review. *Proceedings of the International Astronomical Union*, 9(S310), 194–203.
- Raymond, S. N., O'Brien, D. P., Morbidelli, A., & Kaib, N. A. (2009). Building the terrestrial planets: Constrained accretion in the inner Solar System. *Icarus*, *203*(2), 644–662.

Robertson, M. P., & Joyce, G. F. (2012). The origins of the RNA world. *Cold Spring Harbor Perspectives in Biology*, 4(5), a003608.

Sagan, C. (1980). Cosmos. Random House.

- Satapatha Brahmana (J. Eggeling, Trans.). (1897). https://sacred-texts.com/hin/sbr/index .htm
- Schulze-Makuch, D. (2017). Most Super-Earths Are Likely Not Habitable. In Air and Space Magazine. https://www.airspacemag.com/daily-planet/most-super-earths-are-likely-not-habitable -180965030/
- Segré, D., Ben-Eli, D., Deamer, D. W., & Lancet, D. (2001). The lipid world. *Origins of Life and Evolution of the Biosphere*, *31*(1), 119–145.
- Subramanian, H., Brown, J., & Gatenby, R. (2020). Prebiotic competition and evolution in self-replicating polynucleotides can explain the properties of DNA/RNA in modern living systems. *BMC Evolution-ary Biology*, *20*(1), 1–9.
- The Book of Genesis. (1989). In *New Revised Standard Version Bible*. Division of Christian Education of the National Council of the Churches of Christ in the United States of America.

Totani, T. (2020). Emergence of life in an inflationary universe. Scientific Reports, 10(1), 1–7.

Whitfield, N., & Strong, B. (1966). I Heard It Through the Grapevine.

# CHAPTER 2

# THE PHYLOGENETIC DISTRIBUTION AND FREQUENCY OF SELF-INCOMPATIBILITY IN FABACEAE

Accepted for publication by the International Journal of Plant Sciences on September 22, 2021. See **Appendix A** for publisher's copyright statement.

# 2.1 Abstract

**Premise**. Surprisingly little is known regarding the phylogenetic distribution and frequency of selfincompatibility (SI) in some of the largest angiosperm families, including the legumes (Fabaceae). The ecological and economic significance of the family has motivated an enormous but scattered output of literature concerning its breeding systems and reproductive biology. Since the last synthesis of this literature, we have gained a clearer understanding of the various mechanisms responsible for SI, but data on their number and phylogenetic distribution remains limited.

**Methodology**. We compiled species-level information on legume breeding systems, pollination syndromes, longevity, stature, stem composition, geographic location, and ploidy. The resulting database contains entries for 1167 unique species across 184 genera, accounting for nearly six percent of the family. Using this data, we assess the phylogenetic distribution and frequency of SI in Fabaceae, and evaluate the potential correlation between SI and three other traits: ploidy, stem composition, and longevity. **Pivotal Results**. We estimate that approximately one-fifth of legume species are SI, with the remainder SC. Appropriate controlled pollinations are only available for a small portion of species in our database, and genetic studies are limited to fewer than 0.5% of the family. Moreover, our survey reveals no convincing evidence that a single SI mechanism operates in Fabaceae. Despite uneven coverage and sample of ca. 5% of Fabaceae, the results of our phylogenetic analyses strongly suggest that longevity is statistically significantly correlated with breeding systems while ploidy is not.

**Conclusions**. Our collection reveals great variety of SI expression in the family. Compelling evidence indicates that one of the best-studied SI systems—RNase-based SI—is homologous across core eudicots and ancestral to the group. Previous studies surmise that this system may operate in Fabaceae, but this remains an open question. Whatever its genetic causes among legumes, the reduced seed and fruit set after selfing appears to evolve dynamically and exhibit uncommon variation. We highlight the need for empirical work evaluating SI and outline areas where future efforts may provide disproportionate rewards.

# 2.2 Introduction

Unlike most sexually reproducing groups, the great majority of flowering plants are hermaphrodites. Although they harbor potential for self-fertilization, flowering plants also often maintain mechanisms that reduce selfing or prevent it entirely (Barrett, 1998). Such mechanisms are part of the enormous variation observed in the anatomical and physiological features, collectively known as plant breeding systems, which govern patterns of reproduction in angiosperms (Neal & Anderson, 2005). While some features reduce selfing by morphological or temporal means (e.g., herkogamy or dichogamy), in the present paper we focus on self-incompatibility (SI)—a general term for a suite of postmating, prezygotic, physiological mechanisms that allow a plant to recognize and reject its own pollen. SI is one of the most common and effective means by which higher plants prevent self-fertilization, but unlike features such as herkogamy or dioecy, SI is largely an invisible trait; there is rarely any obvious characteristic morphology associated with it. While all SI systems involve the coordinated action of male and female components facilitating the recognition and rejection of self-pollen, several different mechanisms are known to exist with distinct genetic and molecular underpinnings (Silva & Goring, 2001).

Current estimates indicate that a variety of SI mechanisms evolved at least 35 times, and this is likely a great underestimate (Igić et al., 2008). The underlying molecular machinery generally functions to inhibit germination of self pollen grains on the stigma by preventing elongation of the self pollen tube within the style, often controlled by a single multiallelic locus called the S-locus (de Nettancourt, 2001). Gametophytic SI (GSI) is the most common type of SI and involves the rejection of self pollen as determined by the pollen haplotype. One particular flavor of GSI (RNase-based SI or RSI) is found in several distantly-related core eudicot families—including Cactaceae, Plantaginaceae, Rosaceae, Rubiaceae, Rutaceae, and Solanaceae—with astonishing genetic and molecular similarities, evidence that this system may be homologous and therefore ancestral to core eudicots (Igić & Kohn, 2001; Liang et al., 2020; Ramanauskas & Igić, 2017; Steinbachs & Holsinger, 2002).

The evolution of SI is thought to be associated with increased evolutionary success of angiosperms (Franklin-Tong & Franklin, 2003; Goldberg et al., 2010; Zavada & Taylor, 1986). Traits that influence mating patterns, including genetic relatedness among mated pairs, affect the amount and the distribution of genetic variation in space and over time (Barrett, 2013). Such traits have broad consequences for the evolutionary pathways available to organisms. Thus the presence or absence of SI has two important effects. First, strict outcrossing achieved by SI may commonly increase net diversification rates, the balance of speciation and extinction rates (Goldberg et al., 2010). Second, patterns of genetic variation determine the raw material available for the action of natural selection, so SI ought to also influence the distribution of all other traits.

Indeed, SI has long been observed to causally or indirectly correlate with other traits, including ploidy (Charlesworth, 1985), stem composition (Arroyo, 1981; Barrett, 1998), stature (Scofield & Schultz, 2006), and longevity (Hamrick & Godt, 1996), among others. Associations between breeding system and traits related to lifespan, like stem composition and longevity, are thought to arise partly due to the increased levels of genetic load present in long-lived plant species (Scofield & Schultz, 2006). It is generally observed that woody and perennial species are disproportionately likely to express some form of SI (Barrett, 1998; Friedman, 2020). Annual species, many of which are weedy, instead predominantly self-fertilize, partly due to load dynamics (Lande & Schemske, 1985), but perhaps also providing reproductive assurance to such species during the relatively brief periods when mating is possible (Hamrick & Godt, 1996). Similarly, there appears to be a strong association between increased selfing and polyploidy (Barringer, 2007; Grant, 1956). This association may likewise provide reproductive assurance for recent polyploids that may be reproductively isolated from their diploid progenitors, in addition to reducing the deleterious effects of high genetic load by expanding the number of gene copies available to an individual organism (reviewed in Ramsey & Schemske, 1998). Polyploidy is also known to causally disrupt the molecular mechanism responsible for RSI (Stone, 2002; Stout & Chandler, 1942), and evidence of correlated evolution of self-compatibility (SC) and polyploidy has been observed across groups with RSI (Miller & Venable, 2000; Robertson et al., 2011). Comparative analyses of this kind

are valuable tools that provide insight into the forces shaping trait distribution among angiosperms, but such analyses are not possible without good data.

There are enormous gaps in breeding system data in many families, especially relevant to determining the phylogenetic distribution and frequency of SI across angiosperms. Notably, the molecular mechanism(s) operating in Fabaceae and their phylogenetic distribution within the family remain unknown, despite many years of active literature (e.g., Sims et al., 1999; Aguiar, Vieira, Cunha, & Vieira, 2015; Casey et al., 2010). Fabaceae contains nearly seven percent of all flowering plant species and is abundantly distributed everywhere on Earth, except the extreme polar regions, exhibiting a practically unparalleled diversity of traits (Yahara et al., 2013). The economic and agricultural importance of this group has led to a considerable body of literature regarding their reproductive biology. Arroyo (1981) collated and explored the distribution of breeding systems in Fabaceae with an impressive collection of data containing 364 species, highlighting a general trend of high frequency of SI in woody tropical groups and a low frequency in herbaceous temperate groups. Neither the molecular causes nor confirmation of the distribution of RSI across core eudicots would be uncovered for nearly a decade following her review (McClure et al., 1989), so her paper could not assess information regarding the mechanism responsible for the action of SI in Fabaceae. While characterization of SI systems is still fraught with challenges (e.g., Allen & Hiscock, 2008), recent work takes advantage of the knowledge regarding the distribution of SI and candidate-based sequencing experiments, focused on the metonymous pistil-expressed genes (T2/S-type RNases) in species that possess this system (Liang et al., 2020; Ramanauskas & Igić, 2017, 2021). A nominal sliver of Fabaceae species (six mostly SC species out of ca. 19,500) have previously been unsuccessfully screened for candidate S-RNase genes (Aguiar, Vieira,

Cunha, & Vieira, 2015; Sims et al., 1999). The limited available data on the distribution and phenomenology of SI in the family represents a barrier to a broader understanding the evolutionary causes and consequences of breeding system variation, obscures the mechanisms that may operate, and limits efficient sampling for efforts to uncover the genetic basis of SI in the family.

Here, we aim to examine the distribution and frequency of SI in Fabaceae by recording the currentlyknown literature. We evaluate potentially correlated evolution between breeding system and three traits: ploidy, longevity, and stem composition. Combining the empirical data gathered from pollination experiments, genetic studies, and molecular and physiological work on SI, we assess the overall weight of evidence regarding the claim that RSI may operate in Fabaceae. Finally, we provide a set of recommendations for efficiently uncovering the molecular genetic basis of SI in this and other families.

#### 2.3 Materials and methods

## 2.3.1 Study system

Fabaceae is the third largest family of flowering plants behind Asteraceae and Orchidaceae, and perhaps the most economically important behind Poaceae (Azani et al., 2017). Its species are distributed worldwide and many play an indispensable ecological role as nitrogen fixers. Legumes from each major lineage appear in the fossil record during the middle to late Eocene, suggesting that its six major clades are approximately 50-55 million years old (Lavin et al., 2005).

As currently described, the family contains six subfamilies (Azani et al., 2017). Papilionoideae, the largest and youngest subfamily, contains roughly 14,000 recognized species or approximately 72% of the family. This subfamily is characterized by the highly-specialized papilionoid flower and an increasing trend towards herbaceous growth form, and is also the most widespread of the family. Most agricul-

turally important Fabaceae species belong to this group (e.g., common beans, soy, lentils, peanuts). Caesalpinioideae, which includes Mimosoideae, is a lineage of mostly tropical trees that contains approximately 4,400 species, or 22 percent of the family. The remaining 6% of species belong to the four smallest and oldest subfamilies: Duparquetioideae, Cercidoideae, Detarioideae, and Dialioideae.

#### 2.3.2 Breeding system data

Using the ISI Web of Knowledge Science Citation Index and Google Scholar, we collected published reports with dozens of search terms related to plant mating and breeding systems within Fabaceae. We also included data from conservation reports, local flora and field guides, books, conference presentations, theses, and monographs, and citations recovered within each source. We examined well over 1000 published works. We assembled a primary database from 366, which contained breeding system and associated trait data.

Breeding system data was collected and scored with an accompanying Quality Score (QS) indicator (Grossenbacher et al., 2017). Instances where relative success of selfed and outcrossed handpollinations were reported, or those that directly provided a measure along with transparent calculation procedures, received a score of one (QS=1; highest quality). This high-quality data was used to calculate an index of self-Incompatibility (ISI). Following Lloyd (1968), we define ISI as,

$$ISI = 1 - rac{relative \ selfed \ success}{relative \ outcrossed \ success}$$
,

where relative pollination success is scored as the ratio of fruits (or seeds) set to flowers pollinated. The subtraction from unity differs from Lloyd's formulation, which is an index of self-*compatibility*, and it has the intuitive effect of having the higher ISI denote greater strength of SI (Raduski et al., 2012). When possible, calculation of ISI with fruit set, compared to seed set, is superior because reduced seed set can occur for reasons other than the pre-zygotic action of SI, especially inbreeding depression and resource limitation. Therefore, we used pollination experiments that report relative fruit set (n = 223). When this data was not available, we resorted to calculating ISI with seed set (n = 13). If only manual self-pollinations were performed (without cross-pollinations), and they successfully yielded fruit, these data were assigned a Quality Score of one (QS=1). If hand pollinations were performed but fruit set or seed set was not reported within the study we assigned QS=2, and those that simply stated the breeding system with no pollination experiments reported were given the minimum QS=3. When reports stated the breeding system with an associated citation, which did not address the breeding system of that species, these data were not scored (i.e., "NA"). Our search included papers published through October 2018.

It is possible for ISI to take on negative values in the rare instances (n = 24) when the relative success of self-pollinations is higher than that of cross-pollinations. Within any single population, there is no clear biological basis for consistently increased fruit set following self-pollination versus cross-pollination (Lloyd & Schoen, 1992; Raduski et al., 2012). Instead, negative ISI values likely result from a variety of sources of error including finite sample size, incompatibility of outcross pollen, variance in pollen loads and other aspects of manual pollination, as well as mistakes (Raduski et al., 2012). Consequently, we set all negative values of ISI to zero, which helps facilitate downstream analyses. For reports of a variety of summary statistics, we employed an arbitrary-but-useful approximation: species with ISI > 0.8 were classified as SI (Bawa, 1974; Raduski et al., 2012).

#### 2.3.3 Chromosome counts and ploidy data

When reported alongside breeding system data, information regarding chromosome number and ploidy of investigated species was scored from the original sources. When published reports did not indicate the chromosome numbers or ploidy of the plants tested, such data was instead taken from the Chromosome Counts Database (CCBD; Rice et al., 2015) semi-automatically using the CCDBcurator R package (Rivero et al., 2019). For an initial prune, these records were sorted against a list of all recognized Fabaceae genera from The Plant List (TPL; The Plant List, 2013) yielding a total of 4926 entries. These entries were then formally synonymized using the procedure described below for a total of 4687 entries. When chromosome numbers differed for a single species, the value with the highest number of published reports was chosen (mode, n = 966). When there were an equal number of reports indicating differing chromosome numbers, this data was not used (n = 193). We constrained chromosome numbers, removing entries with extremely low haploid (< 4) values, because the original sources sometimes contain errors because of multivalence or notes of irregular chromosome counts. We observed 17 species (0.3% of the total dataset) with haploid chromosome numbers greater than 56. We cross-referenced these entries with published data to ensure these reports were accurate. Our final dataset of chromosome counts included 4202 unique entries.

**Stebbins' fraction**. Most taxa have chromosome number information but lack ploidy determinations. We inferred species' ploidy designations relative to their generic base with an estimate referred to as the "Stebbins' fraction" (Román-Palacios et al., 2019; Stebbins, 1938). Sporophytic (2n) counts that are greater than, or equal to, 3.5 times the lowest haploid (n) count of the genus are categorized as polyploid (Wood et al., 2009). In addition, we removed genera with data for fewer than three species as the generic base may not be reliably established. While the use of this method to estimate increases in ploidy level may fail to detect older duplication events, we are primarily concerned with uncovering a pattern of association between recent chromosome doubling and loss of SC, indicative of the action of RSI.

## 2.3.4 Other trait data

We collected information on other traits if other traits of interest to use were also concurrently reported in the study with breeding system data for each species. Stem composition, longevity, and stature were almost always included within the original published reports, but in the rare instances they were not, this information was collected from another published report when available.

We also collected data on pollen:ovule ratio (P/O), outcrossing rate, pollination syndromes (including tripping mechanisms, floral morphology that prevents self-fertilization without a pollinator), geographic location, elevation, and pollen limitation (calculated from observations of natural and manual pollination success). This data is not evaluated in the present paper but is available in the associated database.

## 2.3.5 Phylogenetic relationships

We extracted phylogenetic relationships among Fabaceae from a larger dataset used to create the broadly inclusive seed plant phylogeny (Smith & Brown, 2018). Briefly, a sparse multilocus alignment matrix of species across Fabales was constructed using 4,548 taxa from NCBI (retrieved from http://www-personal.umich.edu/~eebsmith/big\_seed\_plant\_datasets/trees/ on March 2, 2018). Tree inference involved constraint (monophyly) of recognized families and a model of sequence evolution partitioned by gene region. We used a list of all recognized Fabaceae genera from TPL to prune

non-legume taxa from the tree. (Taxonomic designations for all species were standardized using the synonym-resolving procedure described below.)

For rate estimates and visualization, we relied on two resulting phylogenetic trees. The first is a species-level tree for trait analysis containing only those species for which our database includes breeding system information and relevant trait data—ploidy, longevity, and stem composition. The second also used the same Fabales phylogeny to construct a genus-level Fabaceae tree by randomly selecting and retaining one tip from each genus. We imposed a semi-parametric penalized likelihood molecular clock (Sanderson, 2002), as implemented by the chronopl function (R package ape, Paradis et al., 2004).

## 2.3.6 Trait correlation analyses

We employed single-variable phylogenetic logistic regression models (Ives & Garland Jr, 2010) to test correlations between breeding system (SI or SC) and three binary traits: ploidy (diploid or polyploid), stem composition (woody or herbaceous), and longevity (annual or perennial). Although the association between annual plants and SC has long been recognized, there is no clear evidence indicating the order of transitions (see Friedman, 2020 for review). Long-lived species (woody or herbaceous perennials) may be associated higher inbreeding depression resulting in selection against selfing, while annual life-cycles present risks without assured reproductive success (Hamrick & Godt, 1996). Based on such observations, Barrett et al. (1996) argue that changes in ecological conditions drive transitions in life history traits, which then have important consequences for breeding system evolution (e.g., the repeated evolution of autogamy within annual lineages of *Gilia* species). Thus, stem composition and longevity were used as explanatory variables, along with ploidy, which commonly disrupts gametophytic SI in diploid taxa (Miller & Venable, 2000; Zenil-Ferguson et al., 2019). Because there are no woody annual species, we could not test interactions between these terms. It is unlikely that longevity and stem composition are independent, and we consequently performed separate analyses with these traits jointly coded as three states: herbaceous annuals, herbaceous perennials, and woody perennials.

We implemented all phylogenetic regression fitting and tests using the function phyloglm in the R package phylolm (Ho et al., 2016). We examined different combinations of these traits to test for significant interactions, but focused only on traits that were significant in single trait models: stem composition and longevity. We also used the jointly coded stem composition and longevity to assess whether or not there may be interactions between these traits and ploidy. We assessed the relative performance of each model using the Akaike's information criterion (AIC) values from the phyloglm results. We display the results of the best performing model using the mosaic function from the R package vcdExtra which also displays Pearson residuals (Friendly, 2021).

## 2.3.7 Taxonomic name resolution

We resolved taxonomic names for all species binomials in breeding system (n = 1330), chromosome count (n = 4957), and phylogenetic datasets (n = 4368) using the Taxonomic Name Resolution Service v5.0 (Boyle et al., 2021) with three sources selected: USDA (National Plant Data Team, 2020), TPL (The Plant List, 2013), and Tropicos (Missouri Botanical Garden, 2020). Our search settings allowed partial matches, and a match accuracy that returned all names found within an edit distance of four (four missed letters, four transposed letters). For all datasets, species designations classified as "Accepted" were kept. Species designations classified as "Synonym", "Illegitimate", "Invalid", "No opinion", "Misapplied", or "Rejected" name were replaced with "Accepted" names when available or removed from the datasets. We also removed species no longer recognized in the family Fabaceae. For detailed synonymy results, see **Appendix A**.

All analyses were conducted in R (R Core Team, 2020). The full breeding system database reference list is available in **Appendix A**.

#### 2.4 Results

The final assembled dataset contains data from 366 reports published from 1886 through 2018, including books, theses, monographs, conservation reports, flora and field guides, and journal papers. There are 1320 entries for 1167 unique currently described species, placed in 184 genera. Our dataset contains 285 entries of data with Quality Score of 1, 104 with Quality Score of 2, and 770 with Quality Score of 3. Of these, 1050 characterize unique currently-described species as SI or SC.

## 2.4.1 Breeding system in Fabaceae

We find that slightly more than 26.4% of tested Fabaceae species are SI and that the highest proportion of SI species appear in Cercidoideae and Detarioideae (Table I). However, only a nominal portion of these subfamilies have been sampled (less than one percent); Caesalpinioideae exhibits the highest proportion of SI species with the highest coverage. The family-wide SI percentage is very close to the 27% SI that Arroyo (1981) originally reported with 364 species. We find the breeding system status for 6.0% of Papilionoideae, 3.5% of Caesalpinioideae, (including 3.6% of Mimosoideae), 3.5% of Dialioideae, less than one percent for Cercidoideae and Detarioideae, and no data for Duparquetioideae, with a total family-wide coverage of 5.4%. We were unable to find any reported data for 410 recognized genera across the family. The vast majority of genera with reported breeding system data have TABLE I: **Summary of SI frequency and breeding system coverage by Fabaceae subfamily.** Columns list the number of total number of unique species, number of species for which data on breeding system is available, the percentage of self-incompatible (SI) species of those tested for breeding system in each subfamily, and coverage, as the percentage of the subfamily with tested breeding system. For comparison with previous studies, the caesalpinioid data originating from the previously-recognized mimiosoid subfamily is shown in italics. Approximate total numbers for each subfamily as given by LPWG Azani et al. (2017).

Subfamily	TOTAL SPECIES	Species w/ data	SI	Coverage
Duparquetioideae	1	0	-	-
Cercidoideae	ca. 335	1	100.0%	0.3%
Detarioideae	ca. 760	3	66.7%	0.4%
Dialioideae	ca. 85	3	33.3%	3.5%
Caesalpinioideae	ca. 4,400	204	52.0%	3.5%
Mimosoideae	ca. 3,000	110	62.7%	3.7%
Papilionoideae	ca. 14,000	839	12.8%	6.0%
Total	ca. <b>19,500</b>	1,050	20.6%	5.4%

such data for less than one-fifth of each genus (**Figure 1**). Of the 163 genera in our dataset that are also found on the phylogenetic tree, 125 are either entirely SI or SC (**Figure 22**, **Appendix A**).

Index of self-incompatibility. We found experimental crossing data sufficient to estimate ISI for 236 species (Figure 2). ISI fruit set data was preferred (n = 223), but we used ISI calculated from seed set where fruit yield was not reported (n = 13). This data spans four subfamilies (Detarioideae, Dialioideae, Caesalpinioideae, and Papilionoideae) and 84 genera. We also added species for which only manual self-pollinations were performed but full fruit set was observed, assigning an ISI value of 0 (n = 66 from Papilionoideae and n = 6 from Caesalpinioideae). Thus, our total ISI dataset contains 308 entries from 95 genera (12.6% of 756 recognized genera). As observed across angiosperms



Figure 1: Genus-level legume phylogeny displaying breeding system data coverage. Each circle is scaled to represent the relative size of genera. Nodes of Caesalpinioideae and Papilionoideae are labeled and illustrations indicate representative morphology of selected genera. Color gradient represents 0-100% breeding system data coverage within that genus. For approximately even coverage, genera were manually binned into each of six percentage-of-coverage categories: 0 (n = 412), > 0 and < 1.7 (n = 31),  $\ge 1.7 and < 4.0 (n = 33)$ ,  $\ge 4.0 and < 9.0 (n = 36)$ ,  $\ge 9.0 and < 20.52 (n = 27)$ , and  $\ge 20.52 (n = 36)$ . The largest category included 412 genera with 0% coverage, indicating that breeding system studies are wanting. (Phylogenetic tree is taken from Smith & Brown, 2018; see text for details).

(Raduski et al., 2012), the frequency appears strongly bimodal, with 79.9% of species possessing values of  $ISI \le 0.2$  or  $ISI \ge 0.8$ . For species where ISI calculation was possible, more than one-third (37.0%) are classified as SI using a cutoff value of 0.8.

We find a significantly higher proportion of SI species in our ISI data (QS1) than across our entire dataset (37.0% versus 20.6%; Two-sample Test for Equality of Proportions,  $\chi^2 = 33.2, p < 0.01$ ). We also observe a higher proportion of SI species for Papilionoideae in our QS1 data than across the dataset (24.8% versus 12.8%; Two-sample Test for Equality of Proportions,  $\chi^2 = 12.2, p < 0.01$ ). This proportional difference in SI data is not significant for Caesalpinioideae (52.0% versus 54.3%; Two-sample Test for Equality of Proportions,  $\chi^2 = 0.1, p = 0.76$ ). (There is a single QS1 record from Detarioideae [SC] and Dialioideae [SI]). This likely reflects a bias caused by the fact that SC is easier to ascertain. It is easier to casually establish self-fertilization, especially in a glasshouse environment, because it requires no specific intervention. On the other hand, it may also often be scored in error, especially when unreliably reported on account of copious fruit set in the field. We ignored such reports, as well as those that claimed entire genera as SI/SC without any accompanying evidence. We cannot be sure, however, that secondary reports did so.

The surveyed literature contains a greater proportion of qualitative SC accounts than qualitative SI accounts (80.0% versus 58.0%; Two-sample Test for Equality of Proportions,  $\chi^2 = 49.8$ , p < 0.01). Of the 900 SC reports, 723 were Quality Score 2 or Quality Score 3 data. Of the 238 SI reports, 138 are Quality Score 2 or Quality Score 3 data. The data also contains a higher proportion of qualitative accounts from Papilionoideae, which has a far lower frequency of SI species generally (84.6% qualitative



Figure 2: Distribution of ISI values in Fabaceae from the collected literature (number of species = 308). ISI measures the relative selfed and outcrossed success of fruit set per flower after manual pollination. ISI = 0 denotes equal success following self- and cross-fertilization, ISI = 1 denotes failed fruit set after self-fertilization, success after cross-fertilization. A critical ISI value of 0.80 is often used to classify species as self-compatible (lower) or self-incompatible (higher).

from Papilionoideae versus 39.8% from other subfamilies; Two-sample Test for Equality of Proportions, $\chi^2 = 181.2, p < 0.01$ ).

**Genetic studies and site of pollen tube inhibition**. We evaluated the evidence consistent with a variety of plausible sterility and SI mechanisms across the family. We found nine studies that investigated the patterns of genetic inheritance of SI in Fabaceae species (Table XI, Appendix A). The results of these studies are sometimes vague, uncertain, difficult to interpret, and may conflate phenomena that are distinct from SI (e.g., inbreeding depression). Only three genera of Papilionoideae have been examined and we found no work establishing the genetics of the response is available for species from other subfamilies.

We also recorded outcomes of studies that tracked site of pollen tube arrest following self-pollination. Published reports on pollen tube arrest were available for 38 species spanning 27 genera two subfamilies: 16 from Papilionoideae, 10 from Caesalpinioideae (five from Mimosoideae), and one from Detarioideae. Fabaceae species show an apparently high variation in the location of pollen tube arrest following self-pollination, which is reported at the stigma, within the style, or at the ovary (**Figure 3**). Half of the studied species are reported to express some kind of "late-acting" phenomenon, meaning that the pollen tube either arrests at the base for the style, at or near the ovary (late-acting SI or LSI), or that the cross fails after fertilization (post-zygotic failure, distinct from SI). The remainder arrest earlier in the pistil, between the stigmatic surface and bottom of the style (Table XII, Appendix A).

# 2.4.2 Trait associations

We evaluated trait correlations between breeding system, ploidy, longevity, and stem composition using phylogenetically-corrected logistic regression (phyloglm function in R). After pruning the



Figure 3: **Site of fertilization failure following self-pollination for the studied Fabaceae species.** For stigma and style, site of self-pollen tube arrest is indicated, and likely involves SI. Species with arrest at the stigma additionally include those where stigmatic membrane disruption can facilitate successful self-pollination (e.g., many *Lotus* species). For ovarian rejection, it is unclear without detailed study whether the failure mechanism is consistent with SI (late-acting SI, or LSI) or whether the failure reflects early-acting inbreeding depression. We include reports of arrest from the nuclear layer of the ovule (*Acacia retinodes*) and any reported "incompatibility" reaction following fertilization (e.g., differential zygote & fruit abortion in selfed species). See **Table XII**, **Appendix A** for a full list of references and study findings.

TABLE II: **Summary of results evaluating trait correlations in Fabaceae**. We performed phylogenetic logistic regression in phyloglm to assess possible correlated evolution between breeding system and three traits: stem composition (StemComp), longevity (Longevity), and ploidy (Ploidy). We also evaluated a collapsed term combining stem composition with longevity (StemCompLong). We report the models that evaluated stem composition and longevity as separate terms, followed by the models that evaluated the jointly-coded stem composition and longevity term (StemCompLong). Significant terms from each model are reported: \*\*\* p < 1e - 06; \*\* p < 0.001; \* p < 0.05; . 0.05 > p < 0.1. Coefficients from each independent term are listed in the order these terms appear in the model. Phylogenetic correlation parameter ( $\alpha$ ) is also reported.

Model	Significant Terms	AIC	Coefficients	α
StemComp + Longevity + Ploidy	Longevity***	140	0.43, 3.79, -0.78	0.81
StemComp + Longevity	Longevity***	142	0.35, 3.53	0.98
Longevity	Longevity***	142	3.65	1.03
StemComp	StemComp*	171	1.56	0.65
Ploidy	None	183	0.22	2.03
StemCompLong + StemCompLong:Ploidy	StemCompLong***,StemCompLong:Ploidy*,Ploidy.	140	2.07, -1.85 (2.10, ploidy)	1.48
StemCompLong + Ploidy + StemCompLong:Ploidy	StemCompLong***,StemCompLong:Ploidy*,Ploidy.	141	2.07, 2.10, -1.85	1.47
StemCompLong	StemCompLong***	145	2.05	1.62
StemCompLong + Ploidy	StemCompLong***	147	2.09, -0.57	1.68

Fabaceae species-level tree to retain only taxa for which ploidy, longevity, and stature were also known, our final tree contained 185 species representing species from Papilionoideae (n = 132), Caesalpinioideae (n = 52), and Cercidoideae (n = 1). This tree and companion dataset were used for all subsequent analyses of trait associations. We report AIC values, significant terms, coefficients of independent variables, and phylogenetic correlation parameters ( $\alpha$ ) for each model (Table II). Our best performing models indicate that longevity is significantly correlated with breeding system.

Stem composition and longevity. Pairwise phylogenetically-corrected logistic regression analyses using the phyloglm function find evidence of correlations between breeding system and both stem composition (z = -1.2, p < 0.01) and longevity (z = -3.3, p < 0.05), but stem composition evaluated independently is one of the lowest performing models examined. Moreover, when stem composition and longevity are evaluated together in a single phyloglm model (StemComp + Longevity), only longevity is significantly correlated with breeding system (z = -3.2; longevity p < 0.05; stem composition p = 0.40). As these terms likely interact, we jointly coded them into a single term Stem-CompLong which was significantly correlated to breeding system in all models evaluated (p < 0.01). The best performing model includes both stem composition and longevity as separate terms, and ploidy, with longevity being the only significant term (z = 4.1, p < 0.01). We observe far fewer woody SC perennials and greater SI perennials regardless of stem composition than expected if these traits were independent (**Figure 4**).

**Ploidy.** Our final chromosome counts database contains 4144 observations across four subfamilies (Cercidoideae, Detarioideae, Caesalpinioideae, Papilionoideae) and 188 genera, with 3294 diploid species and 850 polyploid. The lowest 2n polyploid value is 16 from the genus *Michrocharis*, where 2n = 8 is the generic base. The highest 2n diploid value is 56 from the genus *Leucaena*, where 2n = 112is the highest observed chromosome count in the genus. Of these 4144 observations, 626 overlap with our breeding system database. The highest number of observations come from Papilionoideae (n = 531, ca. 3.8% of the subfamily), followed by Caesalpinioideae (n = 94, ca. 2.1% of the subfamily), and a single observation from Cercidoideae. Of the 541 diploid species, 442 are classified as SC and 99 are SI. Of the 85 polyploid species, 66 are classified as SC and 19 are SI.

In the dataset we used for phylogenetically-corrected logistic regression analysis, we retain 185 species, 33 of which are polyploid. Of the 33 polyploid species, 12 are SI: one from Caesalpinioideae, and 11 from Papilionoideae. If we change the threshold for categorizing polyploids such that sporo-phytic counts (2n) that are greater than or equal to four times the lowest haploid (n) count of the genus



Figure 4: **Mosaic plot for best performing legume trait correlation model.** We selected one of the bestperforming trait correlation models (see Table II for full model results): breeding system as a response variable of stem composition, longevity, and ploidy as separate terms (*StemComp + Longevity + Ploidy*). Individual boxes are sized relative to the number of observations for each group, and numbers in each box display Pearson residuals. Shading reflects the residuals (i.e., lack of fit) for this particular model: for example, we observe far fewer diploid woody perennials than would be expected if these traits were independent. Note that trait evolution of all evaluated models likely violates implicit model assumptions (e.g., equal diversification rates, homogeneity of rates), and these results are highly preliminary. are coded as polyploid, we retain seven SI polyploids: one from Caesalpinioideae, and six from Papilionoideae. We find no evidence of significant correlation between ploidy and breeding system in any models evaluated even with the greater threshold for assigning polyploids, but we do find evidence of a possible interaction between our jointly coded StemCompLong and ploidy (z = 2.6, p < 0.05, Table II).

## 2.5 Discussion

Establishing the pattern of genetic control of SI and gathering functional genetic evidence uncovering the molecular actors is time-consuming and expensive. This partly explains the absence of a clear explanation for the identity of genes governing SI as well as its distribution across legumes. Unlike in Brassicaceae or Solanaceae, for example, where the genetic basis of SI is widely shared by species within each family, it seems that the expression of various phenomena related to SI in Fabaceae is both varied and variable. Despite some lack of clarity, our results confirm many of those recovered by Arroyo (1981) and the greater availability of data, including phylogenetic relationships, allows several insights. SI is broadly distributed in Fabaceae and the variability of its phenomenology, both within and between genera, suggests that no single SI "system" — as they are presently defined and understood — consistently operates or explains this variability. Furthermore, we find no compelling evidence, to date, that the most widespread RNase-based mechanism (RSI) is responsible for causing SI in this family. Such absence of evidence should clearly be interpreted with caution. We report 366 papers with breeding system information, an apparently large number which belies a surprising lack of data on SI frequency, phenomenology, and underlying genetics across this very large family. Below, we discuss the principal findings, some shortcomings of our approaches, and outline areas where future work may be effectively focused.

#### 2.5.1 Features of SI in Fabaceae

Despite long-standing interest in the genetic basis, phylogenetic distribution, and frequency of SI across Fabaceae, the body of literature has not been systematically collected and summarized in decades. We modeled our search on the previous dataset, compiled by Arroyo (1981), which contained information on the breeding systems and their correlates for 364 species. With additional 686 species, we find a similar patterns of SI occurrence across Fabaceae: higher frequency of SI across the primarily woody lineages (Duparquetioideae, Cercidoideae, Detarioideae, Dialioideae, Caesalpinioideae), and a lower frequency in the primarily herbaceous papilionoids (Figure 23, Appendix A).

Arroyo also pointed out that the lower proportion of SI species in Papilionoideae possibly reflects their trend toward herbaceous stem composition. While SI is found in all subfamilies with breeding system data, its occurrence is generally concentrated in some of the less species rich ("early diverging" or "basal") lineages—such as the tribes Bauhineae and Cercideae, as well as some mimosoids. It also appears elsewhere, recently derived within narrowly circumscribed groups, such as the genus *Trifolium*, potentially supporting the claim that some form of SI evolved independently in this family (Arroyo, 1981). Without a clear understanding of the molecular genetic components and broad sampling, however, it is presently impossible to confidently code traits, infer much detail regarding breeding system trait transitions, and or estimate the confounded trait-dependent diversification rates.

Major differences seem to exist in the described phenomenology of SI, both within and between groups. Nevertheless, we recover a bimodal distribution of ISI values, similar to the one described across many distantly-related angiosperm families (e.g., Raduski et al., 2012; Grossenbacher et al., 2017). This widely observed bimodality of ISI may be effected by a combination of population genetic (Porcher & Lande, 2005) and macroevolutionary processes. While SC does not ensure high selfing rates, breeding system (including ISI, its imperfect estimation metric) is strongly associated with mating system (Raduski et al., 2012). Therefore, some of the long-term equilibrium dynamics may be explained by the causal interaction of breeding and mating systems—breakdown of SI allows selfing rates to depart from zero. Species with high ISI values generally have high outcrossing rates while those with low ISI values have lower ones, albeit with a wider range of outcrossing rates, as SC species may employ other characters that promote outcrossing.

The current theoretical approaches fail to accommodate higher-level processes, despite the possibility that transitions may be cladogenetic and highly asymmetric, and differential speciation and extinction rates may be ever present (Goldberg et al., 2010). The expected ISI distribution can then take on a number of shapes, including the observed bimodality, even in the absence of stabilizing lower level selection (Igić & Busch, 2013). It should be noted, however, that intra-specific distributions of ISI also tend to adhere to extreme values, casting considerable doubt on any proposition of persistent stabilizing selection on breeding systems (Raduski et al., 2012).

A challenging feature in determining the distribution and frequency of SI is the wide range in quality of the reported data. Data reporting varies in part because of the varying definitions of SI employed, and methods employed to establish whether or not a plant is SI. We subscribe to the century-old reasoning, which broadly limits SI to pre-fertilization phenomena and considers post-zygotic differences in results of self- and cross-pollinations distinct (perhaps most frequently inbreeding depression, selective abortion, etc.). This definition is consistent with East & Park (1917), Brewbaker (1958), Arasu (1968), and de Nettancourt (2001), which all suppose SI to be the inability of a co-sexual plant producing fertile gametes to set seeds when self-pollinated.

The use of a quality score provides a means of quantifying confidence in the reported data. Manual pollination experiments that allow for the calculation of ISI are the highest quality because these methods distinguish SI from the myriad of other reasons that a plant may not set seed from its own pollen (e.g., resource limitation, unreceptive stigma, immature pollen, or experimental errors). Moreover, when manual self-pollination is performed without manual outcrossing in tandem, this may provide information on the ability to self-fertilize but is uninformative with respect to SI. Yet restricting the survey to only species for which high quality data exists excludes a high proportion of available data. Many informative sources include the incompatibility status of a species without experimental data because that aspect was either not appropriate in that venue (e.g., floras) or was not the purpose of the paper. It is likely that such sources also contain accurate information on SI.

## 2.5.2 Weak evidence regarding mechanism(s) of SI Fabaceae

Several correlated characteristics plausibly implicate RSI in self-rejection within Fabaceae, particularly in *Trifolium* (Casey et al., 2010). These characteristics include the gametophytic SI rejection of pollen established through reciprocal crossing experiments (Brewbaker, 1955), pollen tube arrest in the style following self-pollination (although some evidence implicates stigmatic location; Casey et al., 2010), as well as the presence of wet stigmas and binucleate pollen grains (Brewbaker, 1967). The wide occurrence of RSI across eudicots—especially the close rosid relatives Rosaceae—along with its ancestral condition in core eudicots (Igić & Kohn, 2001) add up to bantamweight of evidence supporting possible occurrence of RSI in Fabaceae. And yet, a closer inspection yields a complicated picture. While most species indeed possess wet styles and binucleate pollen, there is almost no clear shared pattern of SI phenomenology beyond this broad characterization. We found reports of evaluation of the genetic pattern of SI response in eight papilionoid species, half of which are from *Trifolium*. While Casey et al. (2010) located a single S-locus in *Trifolium repens*, their study did not report evidence regarding the presence or absence of candidate T2/S-RNases known to cause the pistil-part response, or SLF genes, which cause pollen-part response. Aguiar, Vieira, Cunha, & Vieira (2015) also screened five papilionoid species for candidate S-RNases and did not find likely actors in SI response, although the majority of their sample was comprised of unambiguously SC species. Finally, RSI mechanism is commonly subject to "competitive interactions," whereby chromosome doubling results in breakdown of SI (Brewbaker, 1954; Lewis, 1947). *T. repens* is a recent tetraploid, which lowers the likelihood that it employs the RSI mechanism (Casey et al., 2010).

Systematic study of the documented site of pollen tube inhibition is informative because provides hints about the possible underlying causes of SI. If pollen rejection occurs on or within the stigmatic depth, it is generally associated with dry stigmatic surface and characterized by SSI genetic mechanisms. Else, if pollen rejection occurs on or within the style, it is generally associated with wet stigmatic surface and characterized by GSI genetic mechanisms (Brewbaker, 1967), which include RSI. Wet stigmatic surfaces are characteristic of Fabaceae (Heslop-Harrison & Shivanna, 1977). And yet, the reported site of inhibition in SI across the family appears unusually varied, with arrest of self-pollen tubes not limited to the style. Pollen tube arrest is noted both on the stigma and in the style, while the majority of reports find that pollen is rejected at the ovule or after fertilization, which indicates phenomena unrelated to

SI.

Angiosperm families were long believed to be generally characterized by a single SI system (Stebbins, 1957), although there is little firm biological basis for such expectations and they were based on sparse empirical data. While evolution of novel self-recognition mechanisms is relatively rare (Igić et al., 2008), the taxonomic rank of "Family" implies no specific age or number of lineages, both of which affect the expected number of novel mechanisms. Primulaceae and Rubiaceae, for example, both harbor both homomorphic GSI and heteromorphic SSI systems (Talavera et al., 2001), and Polemoniaceae may express both GSI and SSI (Goodwillie, 1997). The available data do not support action of a conserved "system" operating in the Fabaceae, or possibly within any one of its genera. It remains to be seen whether any species in Fabaceae express RSI, and our preliminary findings indicate that, should such species exist, they are most likely to be found in the woody species, outside of Papilionoideae.

Finally, species with so-called "late-acting" (LSI) mechanisms—that is, ovarian (ovule penetration) or zygotic (zygote development) arrest—are dramatically understudied across angiosperms, as well as in Fabaceae (Gibbs, 2014). We are not aware of a single functional genetic study uncovering the underlying causes. The task is made more difficult because many studies that encounter a failure of plants to set fruit or seed often casually ascribe such phenomena to LSI, which conflates SI mechanisms of self-rejection with inbreeding depression, resource limitation, and maternal effects. Therefore, the distinction between pre- and post-zygotic arrest may appear merely semantic, but it can bring clarity to functional and mechanistic studies in establishing causality in both applied and basic domains of research (de Nettancourt, 2001). The wide acceptance of a distinction in evolutionary literature likely traces to a rather straightforward conjecture: a chronic wastage of fertilized ovules would be quickly stemmed by natural selection, given any variation in a slightly earlier acting barrier to fertilization. The lack of such a distinction between pre- and post-zygotic phenomena also makes the pursuit of the molecular underpinnings, their evolutionary origins, and mechanisms by which they are maintained (or lost) much more difficult. The disparate nature of the various causes of self-sterility and incompatibility, as well as their inconsistent reporting across the family, all but ensures that they do not describe a single underlying genetic system.

## 2.5.3 Trait associations

Arroyo (1981) noted a correlation between woodiness and SI within Fabaceae. Previous studies have commonly found an association between a perennials and SI (Hamrick & Godt, 1996), as such species may predominantly outcross due to the substantial reduction in fitness among selfed progeny (Morgan et al., 1997). Selection for self-compatibility may also be strong in annual species as a means of reproductive assurance. We conducted analyses that find woodiness and SI may only be correlated through perennial duration, inasmuch as woody species are obligate perennials. In our best performing models, we do not find evidence of a correlation between stem composition (alone) and SI. Instead, our analyses indicate that longevity may be significantly correlated with SI, while stem composition is not.

GSI mechanisms are often causally associated with breakdown of SC following polyploidization, due to S-locus duplication and so-called "competitive interaction" (Lewis, 1947). Several studies have examined the potential correlation between SC and polyploidy. Mable (2004), when considering only groups with GSI, found some evidence of a correlation between polyploidy and SC in a comparative study across angiosperms. Other studies have similarly reported a correlation between SC and polyploidy in families with GSI (Miller & Venable, 2000; Robertson et al., 2011). As such, a lack of correlation between ploidy and breeding system in some legumes discounts the probability of occurrence of RSI. Recent work also supports a long-ago whole genome duplication (WGD) in the common ancestor of papilionoids (Cannon et al., 2015), additional evidence that the presence of RSI is unlikely in this subfamily. Although we find no significant correlation between SC and polyploidy, our analysis was restricted to a small portion of Fabaceae species (approximately 1%), the majority of which are papilionoids. It is as yet unclear if there is a single homologous SI mechanism operating in the family or several independently-evolved mechanisms, so that any interpretations of comparative analyses of this kind are deeply uncertain.

A relatively low frequency of SI and the observed trait associations may have multifarious causes or a single major underlying cause. For example, they may be each be caused by selection imposed by frequent colonization and dispersal mode of speciation. A common parallel outcome of colonization is a filtering effect for uniparental reproduction (Baker, 1955; Grossenbacher et al., 2017). Such a filtering effect may also be amplified in short-lived species (e.g. annuals), given their need for immediate reproduction, and likewise alleviated in long-lived perennials, which would allow more than one immigrant of the same species to accumulate over an extended period of time. Interestingly, although there is a considerable literature on the topic of establishment following long-distance dispersal, especially the role of breeding systems (termed "Baker's rule," reviewed in Pannell et al., 2015) and rhizobial partner nodulation (e.g., Parker, 2001), virtually none of it examines their interaction. While the data on frequency of SI is not particularly reliable, because of its low reporting for most angiosperm families, Igić et al. (2008) surmise it to be ca. 40-50% for flowering plants, overall. The eye-catching low frequency of ca. 20% SI for Fabaceae is also well under the best estimates for Asteraceae (60%), Brassicaceae (30-40%), and Solanaceae (40%) (Grossenbacher et al., 2017).

Phylogenetic comparative analyses, such as the ones we conducted, have some fairly basic limitations. First, even though best evidence strongly links the traits we examined with differential diversification and asymmetric transition rates (Goldberg et al., 2010; Mayrose et al., 2011), because our data are both sparse and unevenly collected, we did not attempt to fit state-dependent diversification models, sometimes referred to as \*SSE family models (Maddison et al., 2007). This is very likely to affect key results. Trait-dependent diversification, of the kind recorded elsewhere with breeding systems, violates the basic assumptions of the widely deployed comparative phylogenetic models (including those we implement in phyloglm). This family of models assumes that traits are essentially passively evolved into a tree of phylogenetic relationships (Maddison, 2006). Violations of state-independent diversification, in turn, interact particularly poorly in case of asymmetric trait changes (Goldberg & Igić, 2008), such as the strongly biased transitions from SI to SC, diploid to polyploid state, woody to herbaceous stem composition, and perennial to annual life history. Second, the haphazard reporting of SI and SC, in particular, and subsequent replication of reporting errors in databases, can clearly affect all downstream comparative analyses, as well as any experimental designs that rely on fidelity in trait reporting. In many ways, the principal value of our paper is the associated database, which attempts to curate the data to both improve accuracy and provide some idea regarding quality assessment for individual data points (using quality scores). Nevertheless, we believe the exploratory analyses to be useful, in the sense that they provide information about the distribution of SI, raw correlations, and summary statistics, each of which can minimally highlight the value of reporting and accounting for breeding system, enabling future assessment of its effect on diversification.

Analyses such as ours are, however, only as good as the data used to conduct them. Recent computational advances enable application of increasingly complex large-scale models, providing the potential to tackle fundamental questions related to the interplay between traits and the mechanisms that shape them, and how this interplay affects the disparity and diversity of organisms. As computational power increases, so must the quality and availability of trait datasets. In Fabaceae, data on SI is absent for over 400 genera, and is generally disproportionately concentrated in groups with agricultural significance. Our (immediately available and unrestricted) comprehensive and curated dataset of breeding systems across Fabaceae can be used to address a variety of linked basic questions in ecology and evolution. The key upshot for work on the molecular genetic basis of breeding systems is that the older caesalpinioid species, with documented pollen tube arrest in the style, may be a good starting point for a candidatebased RNA-seq approach in any future exploration of SI in legumes (as in Ramanauskas & Igić, 2021). SI has extensive and profound consequences in the evolutionary history of angiosperms. Without the continued difficult work of recording and reporting the basic phenomenology of SI, as well as examining its molecular components, it remains an open question whether SI in Fabaceae represents a single system with many losses across groups, or many separate, independently-evolved systems. Meaningful data on SI and its distribution across flowering plants enables determination of the selective pressures influencing both the origin and maintenance of SI systems across macroevolutionary timescales, and helps determine how these systems influence the diversity and ecological success of angiosperms. Such data enables agricultural and conservation efforts and provides improved power for tackling evolutionary problems concerning how reproductive processes shape organisms and their distributions.

#### 2.6 Acknowledgements

We were inspired to pursue this work by M. T. Kalin-Arroyo's paper, published in 1981 without a global electronic data network. We are grateful to the staff at the Richard J. Daley Library, University of Illinois at Chicago, for their help with collecting many difficult-to-find books, papers, and theses. Andrew Raduski and Karolis Ramanauskas provided help and advice throughout the project. The paper was vastly improved by critiques and suggestions from two anonymous reviewers and Editor H. Sauquet. This work was supported by the National Science Foundation grant NSF-DEB-1655692 to B.I.

## 2.7 Cited Literature

- Abberton, M. T. (2007). Interspecific hybridization in the genus *Trifolium*. *Plant Breeding*, *126*(4), 337–342.
- Aguiar, B. (2015). Characterization of the S-locus in Rosidae (Rosaceae and Fabaceae) [PhD Thesis]. Universidade do Porto.
- Aguiar, B., Vieira, J., Cunha, A. E., Fonseca, N. A., Iezzoni, A., van Nocker, S., & Vieira, C. P. (2015). Convergent evolution at the gametophytic self-incompatibility system in *Malus* and *Prunus*. *PloS One*, *10*(5), e0126138.
- Aguiar, B., Vieira, J., Cunha, A. E., & Vieira, C. P. (2015). No evidence for Fabaceae gametophytic self-incompatibility being determined by Rosaceae, Solanaceae, and Plantaginaceae S-RNase lineage genes. *BMC Plant Biology*, *15*(1), 129.
- Allen, A., & Hiscock, S. (2008). Evolution and phylogeny of self-incompatibility systems in angiosperms. In *Self-incompatibility in Flowering Plants* (pp. 73–101). Springer.

Arasu, N. (1968). Self-incompatibility in angiosperms: A review. Genetica, 39(1), 1-24.
- Arroyo, M. K. (1981). Breeding systems and pollination biology in Leguminosae. In R. Polhill & P. Raven (Eds.), *Advances in Legume Systematics, Part 2* (pp. 723–769). Royal Botanic Gardens, Kew.
- Atwood, S. S. (1941). Cytological basis for incompatibility in *Trifolium repens*. American Journal of *Botany*, 28(7), 551–557.
- Atwood, S. S., & Hill, H. D. (1940). The regularity of meiosis in microsporocytes of *Trifolium repens*. *American Journal of Botany*, 730–735.
- Azani, N., Babineau, M., Bailey, C. D., Banks, H., Barbosa, A. R., Pinto, R. B., Boatwright, J. S., Borges,
  L. M., Brown, G. K., Bruneau, A., & others. (2017). A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny: The Legume Phylogeny Working Group (LPWG). *Taxon*, 66(1), 44–77.
- Baker, H. G. (1955). Self-compatibility and establishment after "long-distance" dispersal. *Evolution*, *9*(3), 347–349.
- Barrett, S. C. (1998). The evolution of mating strategies in flowering plants. *Trends in Plant Science*, *3*(9), 335–341.
- Barrett, S. C. (2013). The evolution of plant reproductive systems: How often are transitions irreversible? *Proceedings of the Royal Society B: Biological Sciences*, *280*(1765), 20130913.
- Barrett, S. C. H., Harder, L. D., & Worley, A. C. (1996). The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1345), 1271–1280.
- Barringer, B. C. (2007). Polyploidy and self-fertilization in flowering plants. *American Journal of Botany*, *94*(9), 1527–1533.

- Bawa, K. S. (1974). Breeding systems of tree species of a lowland tropical community. *Evolution*, 28(1), 85–92.
- Boyle, B. L., Matasci, N., Mozzherin, D., Rees, T., Barbosa, G. C., Kumar Sajja, R., & Enquist, B. J. (2021). Taxonomic name resolution service, version 5.0. *Botanical Information and Ecology Network*. https://tnrs.biendata.org/
- Brewbaker, J. L. (1954). Incompatibility in autotetraploid *Trifolium repens* L. I. Competition and self-compatibility. *Genetics*, *39*(3), 307.
- Brewbaker, J. L. (1955). Studies of oppositional allelism in *Trifolium nigrescens*. *Hereditas*, *41*(3-4), 367–375.
- Brewbaker, J. L. (1958). Self-compatibility in tetraploid strains of Trifolium hybridum. *Hereditas*, 44(4), 547–553.
- Brewbaker, J. L. (1967). The distribution and phylogenetic significance of binucleate and trinucleate pollen grains in the angiosperms. *American Journal of Botany*, *54*(9), 1069–1083.
- Brink, R., & Cooper, D. (1939). Somatoplastic sterility in Medicago sativa. Science, 90(2345), 545-546.
- Bruneau, A., & Anderson, G. J. (1988). Reproductive biology of diploid and triploid *Apios americana* (Leguminosae). *American Journal of Botany*, *75*(12), 1876–1883.
- Cannon, S. B., McKain, M. R., Harkess, A., Nelson, M. N., Dash, S., Deyholos, M. K., Peng, Y., Joyce, B., Stewart Jr, C. N., Rolf, M., & others. (2015). Multiple polyploidy events in the early radiation of nodulating and nonnodulating legumes. *Molecular Biology and Evolution*, *32*(1), 193–210.

- Casey, N. M., Milbourne, D., Barth, S., Febrer, M., Jenkins, G., Abberton, M. T., Jones, C., & Thorogood,
  D. (2010). The genetic location of the self-incompatibility locus in white clover (*Trifolium repens*L.). *Theoretical and Applied Genetics*, *121*(3), 567–576.
- Charlesworth, D. (1985). Distribution of dioecy and self-incompatibility in angiosperms. In *Evolution: Essays in Honour of John Maynard Smith* (pp. 237–268). Cambridge University Press.
- Charlesworth, D. (2006). Evolution of plant breeding systems. *Current Biology*, 16(17), R726–R735.
- de Nettancourt, D. (1977). Incompatibility in Angiosperms (1st ed., Vol. 3). Springer.
- de Nettancourt, D. (1999). Homomorphic Self-Incompatibility in Flowering Plants. In *Fertilization in Higher Plants* (pp. 101–112). Springer.
- de Nettancourt, D. (2001). Incompatibility and Incongruity in Wild and Cultivated Plants (Second). Springer-Verlag.
- Dobrofsky, S., & Grant, W. (1980). Electrophoretic evidence supporting self-incompatibility in *Lotus corniculatus*. *Canadian Journal of Botany*, *58*(6), 712–716.
- Doyle, J. J. (2012). Polyploidy in Legumes. In Polyploidy and Genome Evolution (pp. 147–180). Springer.
- East, E. M., & Park, J. B. (1917). Studies on self-sterility I. The behavior of self-sterile plants. *Genetics*, 2(6), 505.
- Franklin-Tong, V. E., & Franklin, F. (2003). The different mechanisms of gametophytic selfIncompatibility. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1434), 1025–1032.
- Friedman, J. (2020). The evolution of annual and perennial plant life histories: Ecological correlates and genetic mechanisms. *Annual Review of Ecology, Evolution, and Systematics*, *51*, 461–481.

Friendly, M. (2021). vcdExtra. https://CRAN.R-project.org/package=vcdExtra

- Fujii, S., Kubo, K.-i., & Takayama, S. (2016). Non-self-and self-recognition models in plant self-incompatibility. *Nature Plants*, *2*(9), 1–9.
- Gibbs, P. E. (2014). Late-acting self-incompatibilitythe pariah breeding system in flowering plants. *New Phytologist*, *203*(3), 717–734.
- Giles, W. L. (1949). *The morphological aspects of self-sterility in* Lotus corniculatus *L*. [PhD Thesis]. University of Missouri.
- Goldberg, E. E., & Igić, B. (2008). On phylogenetic tests of irreversible evolution. *Evolution*, 62(11), 2727–2741.
- Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., & Igić, B. (2010). Species selection maintains self-incompatibility. *Science*, *330*(6003), 493–495.
- Goodwillie, C. (1997). The genetic control of self-incompatibility in *Linanthus Parviflorus* (Polemoniaceae). *Heredity*, *79*(4), 424–432.
- Grant, V. (1956). The influence of breeding habit on the outcome of natural hybridization in plants. *The American Naturalist*, *90*(854), 319–322.
- Grossenbacher, D. L., Brandvain, Y., Auld, J. R., Burd, M., Cheptou, P.-O., Conner, J. K., Grant, A. G., Hovick, S. M., Pannell, J. R., Pauw, A., & others. (2017). Self-compatibility is over-represented on islands. *New Phytologist*, *215*(1), 469–478.
- Hamrick, J., & Godt, M. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1345), 1291–1298.

- Hauck, N. R., Yamane, H., Tao, R., & Iezzoni, A. F. (2002). Self-compatibility and incompatibility in tetraploid sour cherry (*Prunus cerasus* L.). *Sexual Plant Reproduction*, *15*(1), 39–46.
- Hauck, N. R., Yamane, H., Tao, R., & Iezzoni, A. F. (2006). Accumulation of nonfunctional S-haplotypes results in the breakdown of gametophytic self-incompatibility in tetraploid *Prunus*. *Genetics*, *172*(2), 1191–1198.
- Heslop-Harrison, J., & Heslop-Harrison, Y. (1982). Pollen-stigma interaction in the Leguminosae: Constituents of the stylar fluid and stigma secretion of *Trifolium pratense* L. *Annals of Botany*, 49(6), 729–735.
- Heslop-Harrison, Y., & Shivanna, K. (1977). The receptive surface of the angiosperm stigma. *Annals of Botany*, *41*(6), 1233–1258.
- Ho, L. S. T., Ane, C., Lachlan, R., Tarpinian, K., Feldman, R., Yu, Q., van der Bijl, W., Maspons, J., Vos, R.,
  & Ho, M. L. S. T. (2016). Package "phylolm". See Http://Cran. R-Project. Org/Web/Packages/Phylolm/Index.
  Html (Accessed February 2018).
- Holsinger, K. E. (2000). Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences*, 97(13), 7037–7042.
- Horovitz, A., Meiri, L., & Beiles, A. (1976). Effects of ovule positions in Fabaceous flowers on seed set and outcrossing rates. *Botanical Gazette*, *137*(3), 250–254.
- Igić, B., & Busch, J. W. (2013). Is self-fertilization an evolutionary dead end? *New Phytologist*, 198(2), 386–397.
- Igić, B., & Kohn, J. R. (2001). Evolutionary relationships among self-incompatibility RNases. *Proceedings of the National Academy of Sciences*, *98*(23), 13167–13171.

- Igić, B., Lande, R., & Kohn, J. R. (2008). Loss of self-incompatibility and its evolutionary consequences. *International Journal of Plant Sciences*, *169*(1), 93–104.
- Ives, A. R. (2019). R<sup>2</sup>s for correlated data: Phylogenetic models, LMMs, and GLMMs. *Systematic Biology*, 68(2), 234–251.
- Ives, A. R., & Garland Jr, T. (2010). Phylogenetic logistic regression for binary dependent variables. *Systematic Biology*, *59*(1), 9–26.

Ives, A. R., & Li, D. (2018). Rr2 (R package). https://github.com/arives/rr2

- Kalisz, S., Ree, R. H., & Sargent, R. D. (2006). Linking floral symmetry genes to breeding system evolution. *Trends in Plant Science*, *11*(12), 568–573.
- Kenrick, J., & Knox, R. B. (1989). Quantitative analysis of self-incompatibility in trees of seven species of *Acacia*. *Journal of Heredity*, *80*(3), 240–245.
- Kubo, K.-i., Paape, T., Hatakeyama, M., Entani, T., Takara, A., Kajihara, K., Tsukahara, M., Shimizu-Inatsugi, R., Shimizu, K. K., & Takayama, S. (2015). Gene duplication and genetic exchange drive the evolution of S-RNase-based self-incompatibility in *Petunia*. *Nature Plants*, *1*(1), 1–9.
- Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, *34*(7), 1812–1819.
- Lande, R., & Schemske, D. W. (1985). The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution*, *39*(1), 24–40.
- Lavin, M., Herendeen, P. S., & Wojciechowski, M. F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology*, *54*(4), 575–594.
  Lewis, D. (1944). Incompatibility in plants. *Nature*, *153*, 575–578.

- Lewis, D. (1947). Competition and dominance of incompatibility alleles in diploid pollen. *Heredity*, *1*(1), 85–108.
- Lewis, D., & Modlibowska, I. (1942). Genetical studies in pears. IV. Pollen-tube growth and incompatibility. *Journal of Genetics, xliii,* 211–222.
- Li, D., Trotta, L., Marx, H. E., Allen, J. M., Sun, M., Soltis, D. E., Soltis, P. S., Guralnick, R. P., & Baiser,
  B. (2019). For common community phylogenetic analyses, go ahead and use synthesis phylogenies. *Ecology*, 100(9), e02788.
- Liang, M., Cao, Z., Zhu, A., Liu, Y., Tao, M., Yang, H., Xu, Q., Wang, S., Liu, J., Li, Y., & others. (2020). Evolution of self-compatibility by a mutant S m-RNase in citrus. *Nature Plants*, 6(2), 131–142.
- Lloyd, D. (1968). Pollen tube growth and seed set in self-incompatible and self-compatible *Leavenworthia* (Cruciferae) populations. *New Phytologist*, *67*(1), 179–195.
- Lloyd, D. G. (1965). Evolution of self-compatibility and racial differentiation in *Leavenworthia* (Cruciferae). In R. Rollins & R. Foster (Eds.), *Contributions from the Gray Herbarium of Harvard University* (Vol. 195, pp. 3–134). The Grey Herbarium of Harvard University.
- Lloyd, D. G., & Schoen, D. J. (1992). Self-and cross-fertilization in plants. I. Functional dimensions. International Journal of Plant Sciences, 153(3, Part 1), 358–369.
- Lundqvist, A. (1993). The self-incompatibility system in *Lotus tenuis* (Fabaceae). *Hereditas*, 119(1), 59–66.
- Mable, B. K. (2004). Polyploidy and self-compatibility: Is there an association? *New Phytologist*, *162*(3), 803–811.

- Maddison, W. P. (2006). Confounding asymmetries in evolutionary diversification and character change. *Evolution*, 60(8), 1743–1746.
- Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, *56*(5), 701–710.
- Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2011). Recently formed polyploid plants diversify at lower rates. *Science*, *333*(6047), 1257–1257.
- Mazerolle, M. J. (2020). AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c) [Manual]. https://cran.r-project.org/package=AICcmodavg
- McClure, B. (2009). Darwin's foundation for investigating self-incompatibility and the progress toward a physiological model for S-RNase-based SI. *Journal of Experimental Botany*, *60*(4), 1069–1081.
- McClure, B. A., Haring, V., Ebert, P. R., Anderson, M. A., Simpson, R. J., Sakiyama, F., & Clarke, A.
   E. (1989). Style self-incompatibility gene products of *Nicotlana alata* are ribonucleases. *Nature*, 342(6252), 955–957.
- Miller, J., & Venable, D. L. (2000). Polyploidy and the evolution of gender dimorphism in plants. *Science*, *289*(5488), 2335–2338.
- Miri, R., & Bubar, J. (1966). Self-incompatibility as an outcrossing mechanism in birdsfoot trefoil (*Lotus corniculatus*). *Canadian Journal of Plant Science*, *46*(4), 411–418.

Missouri Botanical Garden. (2020). Tropicos. http://www.tropicos.org

Morgan, M. T., Schoen, D. J., & Bataillon, T. M. (1997). The evolution of self-fertilization in perennials. *The American Naturalist*, 150(5), 618–638.

National Plant Data Team. (2020). USDA, NRCS: The PLANTS Database. http://plants.usda.gov

- Neal, P., & Anderson, G. (2005). Are 'mating systems' "breeding systems" of inconsistent and confusing terminology in plant reproductive biology? Or is it the other way around? *Plant Systematics and Evolution*, *250*(3), 173–185.
- Pannell, J. R., Auld, J. R., Brandvain, Y., Burd, M., Busch, J. W., Cheptou, P.-O., Conner, J. K., Goldberg, E. E., Grant, A.-G., Grossenbacher, D. L., & others. (2015). The scope of Baker's law. *New Phytologist*, *208*(3), 656–667.
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*(2), 289–290.
- Parker, M. A. (2001). Mutualism as a constraint on invasion success for legumes and rhizobia. *Diversity and Distributions*, *7*(3), 125–136.
- Porcher, E., & Lande, R. (2005). Loss of gametophytic self-incompatibility with evolution of inbreeding depression. *Evolution*, *59*(1), 46–60.
- Raduski, A. R., Haney, E. B., & Igić, B. (2012). The expression of self-incompatibility in angiosperms is bimodal. *Evolution*, 66(4), 1275–1283.
- Ramanauskas, K., & Igić, B. (2017). The evolutionary history of plant T2/S-type ribonucleases. *PeerJ*, *5*, e3790.
- Ramanauskas, K., & Igić, B. (2021). RNase-based self-incompatibility in cacti. *New Phytologist, 231*(5), 2039–2049.
- Ramsey, J., & Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, *29*(1), 467–501.

- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
- Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O.,
  & Mayrose, I. (2015). The chromosome counts database (CCDB) a community resource of plant
  chromosome numbers. *New Phytologist*, 206(1), 19–26.
- Rivero, R., Sessa, E. B., & Zenil-Ferguson, R. (2019). EyeChrom and CCDB curator: Visualizing chromosome count data from plants. *Applications in Plant Sciences*, *7*(1), e01207.
- Robertson, K., Goldberg, E. E., & Igić, B. (2011). Comparative evidence for the correlated evolution of polyploidy and self-compatibility in Solanaceae. *Evolution: International Journal of Organic Evolution*, 65(1), 139–155.
- Rodriguez-Riano, T., Ortega-Olivencia, A., & Devesa, J. (1999). Reproductive biology in two *Genisteae* (Papilionoideae) endemic of the western Mediterranean region: *Cytisus striatus* and *Retama sphae-rocarpa*. *Canadian Journal of Botany*, *77*(6), 809–820.
- Román-Palacios, C., Molina-Henao, Y. F., & Barker, M. S. (2019). Polyploidy increases overall diversity despite higher turnover than diploids in the Brassicaceae. *bioRxiv*, 717306.
- Sandal, P. (1951). *Mechanisms of self- and cross-incompatibility in sweet clover,* Melilotus officianalis *Lam.* [PhD Thesis]. Iowa State College.
- Sanderson, M. J. (2002). Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution*, *19*(1), 101–109.

- Sassa, H., Nishio, T., Kowyama, Y., Hirano, H., Koba, T., & Ikehashi, H. (1996). Self-incompatibility (S) alleles of the Rosaceae encode members of a distinct class of the T-2/S ribonuclease superfamily. *Molecular and General Genetics MGG*, *250*(5), 547–557.
- Schemske, D. W., & Lande, R. (1985). The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution*, *39*(1), 41–52.
- Scofield, D. G., & Schultz, S. T. (2006). Mitosis, stature and evolution of plant mating systems: Low- $\phi$  and high- $\phi$  plants. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1584), 275–282.
- Seavey, S. R., & Bawa, K. S. (1986). Late-acting self-incompatibility in angiosperms. *The Botanical Review*, *52*(2), 195–219.
- Silva, N., & Goring, D. (2001). Mechanisms of self-incompatibility in flowering plants. *Cellular and Molecular Life Sciences CMLS*, 58(14), 1988–2007.
- Sims, T., Ordanic, M., Fjellstrom, R., & Collins, P. (1999). Molecular genetics of gametophytic selfincompatibility in *Petunia hybrida* and *Trifolium pratense*. (Abstract only). *Pollen-Stigma Interactions*, 52.
- Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. American Journal of Botany, 105(3), 302–314.
- Stebbins, G. L. (1938). Cytological characteristics associated with the different growth habits in the dicotyledons. *American Journal of Botany*, 189–198.
- Stebbins, G. L. (1957). Self fertilization and population variability in the higher plants. *The American Naturalist*, *91*(861), 337–354.

- Steinbachs, J., & Holsinger, K. (2002). S-RNase mediated gametophytic self-incompatibility is ancestral in Eudicots. *Molecular Biology and Evolution*, *19*(6), 825–829.
- Stone, J. (2002). Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. *The Quarterly Review of Biology*, *77*(1), 17–32.
- Stout, A. B., & Chandler, C. (1942). Hereditary transmission of induced tetraploidy and compatibility in fertilization. *Science*, *96*(2489), 257–258.
- Talavera, S., Gibbs, P. E., Fernaández-Piedra, M. P., & Ortiz-Herrera, M. A. (2001). Genetic control of self-incompatibility in *Anagallis monelli* (Primulaceae: Myrsinaceae). *Heredity*, *87*(5), 589–597.
- Tandon, R., Shivanna, K., & Ram, H. (2001). Pollination biology and breeding system of Acacia senegal. Botanical Journal of the Linnean Society, 135(3), 251–262.

The Plant List. (2013). The Plant List (Nos. v1.1). https://www.theplantlist.org/

- Weller, S. (1995). The evolution of self-incompatibility in flowering plants: A phylogenetic approach. *Experimental and Molecular Approaches to Plant Biosystematics*, 53, 355–382.
- Whitehouse, H. L. (1950). Multiple-allelomorph incompatibility of pollen and style in the evolution of the angiosperms. *Annals of Botany*, *14*(54), 199–216.
- Williams, J. S., Wu, L., Li, S., Sun, P., & Kao, T.-H. (2015). Insight into S-RNase-based self-incompatibility in *Petunia*: Recent findings and future directions. *Frontiers in Plant Science*, 6(41), 1–6.
- Williams, W. (1951). Genetics of incompatibility in alsike clover, *Trifolium hybridum*. *Heredity*, *5*(1), 51–73.

- Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009).
   The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences*, *106*(33), 13875–13879.
- Xue, Y., Carpenter, R., Dickinson, H. G., & Coen, E. S. (1996). Origin of allelic diversity in *antirrhinum* S locus RNases. *The Plant Cell*, 8(5), 805–814.
- Yahara, T., Javadi, F., Onoda, Y., de Queiroz, L. P., Faith, D. P., Prado, D. E., Akasaka, M., Kadoya, T., Ishihama, F., Davies, S., & others. (2013). Global legume diversity assessment: Concepts, key indicators, and strategies. *Taxon*, *62*(2), 249–266.
- Yang, X., Wang, Y., Si, X., & Feng, G. (2020). Species traits linked with range shifts of Chinese birds. *Global Ecology and Conservation*, *21*, e00874.
- Zavada, M. S., & Taylor, T. N. (1986). The role of self-incompatibility and sexual selection in the gymnosperm-angiosperm transition: A hypothesis. *The American Naturalist*, *128*(4), 538–550.
- Zenil-Ferguson, R. (2018). Clean angiosperm records.
- Zenil-Ferguson, R., Burleigh, J. G., Freyman, W. A., Igić, B., Mayrose, I., & Goldberg, E. E. (2019). Interaction among ploidy, breeding system and lineage diversification. *New Phytologist*, 224(3), 1252– 1265.
- Zenil-Ferguson, R., Ponciano, J. M., & Burleigh, J. G. (2017). Testing the association of phenotypes with polyploidy: An example using herbaceous and woody eudicots. *Evolution*, *71*(5), 1138–1148.

# **CHAPTER 3**

### THE ORCHIDS AND THEIR BREEDING SYSTEMS

Perhaps the most enduring question in evolutionary biology is why some groups of organisms are more diverse than others. There are roughly 6,000 species of mammals, including us *Homo sapiens* but more than 28,000 species of orchids found across five subfamilies, not including cultivar or hybrid varieties. In fact, there are more orchid species than all mammals, birds, and reptiles combined (Givnish et al., 2015). Why? One possible explanation is that members of the orchid family (Orchidaceae) appeared earlier in evolutionary history, and that the size of the family reflects a longer period of time on Earth. More time on Earth could mean more time to diversify. But as it turns out, the opposite is true: the common ancestor of mammals pre-dates the earliest orchid species by roughly 50 million years (Poinar & Rasmussen, 2017). Moreover, orchid species are found everywhere on the planet, from the arctic to the tropics, exhibiting an exceptional variety of highly-specialized floral morphology, pollinator relationships, and growth form. So then, what does account for this incredible diversity?

Three years after the publication of his seminal work *On the Origin of Species by Means of Natural Selection*, Charles Darwin published his massive treatise on orchids (Darwin, 1862). In seeking to bolster evidence for his claims in *Origin* and demonstrate his case for natural selection, he painstakingly examined the floral morphology and pollination biology of orchid species from more than 63 genera (Yam et al., 2009). (In later editions, this was expanded to include 85 orchid genera.) With *Orchids*, Darwin provided early published accounts of what would later collectively come to be known as *plant breeding systems*, the anatomical and physiological traits that govern how plants mate. He also began laying an empirical foundation for one of the central themes woven throughout his work (Harder & Johnson, 2009): in nature, he argued, cross-fertilization is dramatically favored over self-fertilization.

To understand this claim, it helps to know something about how orchids mate, especially the structure and function of the orchid's main showpiece and reproductive organ, the flower. Sex itself is essentially as old as dirt—originating more than a billion years ago—and it's also widespread and remarkably similar across the tree of life, from fungi to fishes: the process almost always involves the fertilization of large, protein-rich gametes by smaller motile gametes. (Biologists generally define the larger gametes as "female" and the smaller gametes as "male.") Although sexual function can be split up in many different ways, including separate male and female organisms as in mammals, nearly all orchids are hermaphrodites, with flowers that produce *both* male gametes (pollen, which orchids package in sugary sacs called *pollinia*) and female gametes (*ovules*). Does this mean that orchids can fertilize their own ovules? In fact, many can. Plants that can successfully produce seed from their own pollen are *self-compatible* (SC). Yet there are many orchids— and estimates suggest more than half of *all* flowering plant species—that possess some mechanism that reduces self-fertilization or prevents it entirely. That is to say, such mechanisms favor cross-fertilization over self-fertilization as Darwin noted in *Orchids* more than 150 years ago.

For two decades prior to the publication of *Orchids*, and for another two after, Darwin was fascinated by the various floral adaptations promoting cross-fertilization—"contrivances," as he called them (Darwin, 1872). But no contrivances fascinated him more than those found in orchids. "Orchids," he wrote to a friend in 1861, "have interested me more than almost anything in my life" (Darwin, 1861b). A significant claim for a man known to perform exhaustive and unusual experiments in his relentless



Figure 5: **The orchid flower.** (a) The orchid's column, which houses the male and female floral components. The anther produces the male gametes, pollinia, while the ovules are located at the base of the column itself. Pollen grains must travel down the length of the style to reach the ovules. (b) The orchid's flower, including the "lip petal" or labellum, which provides a path for insects to reach the entrance to the column at the center of the flower. (Digital floral illustration adapted from Linden, 1885; digital column illustration inspired by Aceto & Gaudio, 2011.)

quest to gather evidence for his theory, even submerging the amputated legs of dead ducks in tubs brimming with snails to determine how snails might disperse across oceans (Quammen, 2007). But for Darwin, orchids were, above all, the quintessential example of his claim that "nature... abhors perpetual self-fertilization" (Darwin, 1862, pg. 359)—and orchid species embodied this claim by use of a stunning suite of baroque adaptations. Many of these adaptations work by separating male and female floral components in space or in time. To make seed, pollen grains must first reach the *stigma*—from there the grains elongate, forming protracted tubes that travel down the length of the *style* until finally reaching ovules located at the base of the flower (**Figure 5**). *Herkogamy* involves specialized floral morphology such that the pollen-producing components (*anthers*) are unable to touch the stigma without the aid of a pollinator. Herkogamy is known to occur in hundreds of orchid species that are otherwise SC, increasing the odds of cross-fertilization facilitated by insects or animals. Indeed, most orchid species employ some variation on a "lip petal," also known as a *labellum*—a tongue-like structure that serves as a landing pad for insects, guiding them down the path of pollination. But in the yellow lady's slipper (*Cyprepedium calceolus*), the labellum resembles more of an oblong bowl (or, more traditionally, the toe of a slipper) than a landing pad (**Figure 6**). Darwin wondered why. He carefully probed the two smaller openings at the top of the flower, assuming that an insect would land on the edge of the bowl, noting the ease with which sticky pollinia adhered to the bristle of his brush. But something puzzled him: sure, the pollinia stuck to his bristle, but positioned on the edge of the bowl it would be nearly impossible for an insect to reach the stigma. How would the thing get fertilized?

On the advice of his friend Asa Gray, an American botanist, he tried depositing a small bee into the bowl's "chamber" and observed something extraordinary (Alcock, 2006). The smooth walls of the chamber made it impossible for the bee to escape—like a fly smacking itself against the glass of a window, the bee slipped and fell back every time it tried to exit out of the larger opening. To escape, the bee was forced upwards along the orchid's *column*—the structure housing the male and female organs and escaped via one of the smaller openings at the top of the flower that Darwin had initially probed



Figure 6: **Botanical illustration of the yellow lady's slipper.** (a) The characteristic flower with the modified labellum cup for "catching" bees. (b) A sideways view of the flower: the bee must crawl out through the top, passing the column along the way. (c) The stigma and anther cap. (d) A cross-section of the ovules. (Digital illustration adapted from Correvon, 1899.)

with bristles. Burrowing out of the small passage covered the bee in gummy pollinia. Upon dissecting the flower, Darwin found that the column was structured so that the bee *first* passed the stigma, and only then became covered. He realized that this morphological feature would make it nearly impossible for a bee to pollinate a flower with its own pollinia. To pollinate the lady slipper, the bee must already be covered in pollinia before entering the chamber and burrowing its way out. The bee would subsequently pick up fresh pollinia, beginning the process once more. A marvelous contrivance—and yet, the yellow lady slipper is only one of many increasingly tricky floral contraptions Darwin recorded while collecting data for *Orchids*.

In the Catasetinae—a group Darwin referred to as "the most remarkable of all orchids" (Pérez-Escobar et al., 2016)—many species separate the maturation of male and female components in time, meaning that the stigma is not receptive when pollinia are produced by the anthers (called *dichogamy*). Some plants even segregate male and female components into entirely separate unisexual flowers (*monoecy*) or into separate unisexual plants (*dioecy*). Darwin was the first to identify these as unisexual dioecious plants in *Catasetum* (Figure 7); before his discovery the male and female plants were actually classified as separate species of entirely different genera, so radical their differences in appearance (Alcock, 2006). Observed in nature (Romero & Nelson, 1986), *Catasetum* species are pollinated exclusively by euglossine bees that covet the chemical attractants produced by these flowers as lavish perfumes for potential mates. As bees collect from male flowers, grabbing chemicals with the front legs and transferring them to the hind legs, they eventually trigger the explosive slingshot release of the anther cap containing the pollinia. The bees' tiny bodies are slapped so aggressively with pollinia that they cease to visit these male flowers again. (The aversion is well-deserved: in one instance, a bee was observed to lose its life



Figure 7: **Botanical illustration of dioecous** *Catasetum discolor*, showing the (a) male and (b) female floral varieties. (Digital illustration adapted from Linden, 1885a.)

to this mechanism.) Instead, beaten-up bees turn to female flowers to collect their perfumes. Female flowers are inverted in shape, allowing the bee to easily deposit the pollinia on its back onto the stigma. This violent pollination mechanism is hypothesized to promote successful fertilization by exploiting bees' aversion to male flowers as soon as they become whacked—sending them directly to the gentler female flowers that are perfectly positioned to accept the fresh packages of pollinia.

Such morphological and temporal adaptations reduce self-fertilization but do not necessarily prevent it entirely. As it turns out, Darwin scratched the surface of something in *Orchids* that would ultimately fascinate biologists for the next century and beyond—outside of these trickier floral contraptions, some plants actually possess invisible *genetic* mechanisms that, once fertilized, allow them to recognize their own pollen and reject it. Known under the broad umbrella term *self-incompatibility* (SI), there is no characteristic morphology associated with these genetic mechanisms; you can't just examine a plant and determine if it is SI. Perhaps this is why Darwin, responsible for arguably the single most important contribution to the field of biology but who knew nothing of genes, danced around the discovery of this trait without ever figuring it out. He wrote, "If we now turn to the more immediate cause of self-sterility [SI], we clearly see that in most cases it is determined by the conditions to which the plants have been subjected" (Darwin, 1872, pg 342). What he meant was that this trait was likely the cause of environmental factors, and not something innate in the plant itself, written in its DNA. He was wrong.

It has taken nearly a century of empirical work to advance our knowledge of SI, and there are still many unknowns. We do know that there exist a great variety of SI systems, widely distributed across all flowering plants (Allen & Hiscock, 2008). Some of the most well-studied systems involve the in-

teraction of an enormous number of linked genes and complex molecular components that rival the immune system in humans. In fact, many SI systems involve so many codependent genes that random mutation between generations can easily cause the system to break, producing a newly-SC plant from an SI parent. These observations present a great many puzzles to contemporary biologists: how do such complex and coordinated genes evolve in the first place? What is so special about these systems that they have evolved independently in many lineages, at least 35 different times (lgić et al., 2008)? How many different SI systems are out there, really? And finally, if mutation can so easily cause such systems to break down, why have they persisted for millions of years?

The first step in answering such questions is determining how SI is distributed across the roughly 400,000 species of flowering plants—that is to say, which species are SI and which are not. Even Darwin, who was largely misguided on SI, understood the complexity of this task. For starters, distinguishing an SI plant involves detailed crossing experiments that are especially challenging in orchids. A great variety of unrelated plants are required, and few collections contain the number of necessary specimens. Second, determining the *expression* of SI in a plant—how SI works, what molecular mechanism may be responsible for it, and its genetic basis—is even more challenging. Many orchids have enormously long *generation times* (the average time from seed production in the parent to seed production in the offspring), some more than a decade, meaning experiments of this kind would outlast the humans performing them. This may be why information on SI in orchids remains limited (Johansen, 1990).

Yet SI is still known to exist in orchids. The most famous example occurs in one of the largest genera, *Dendrobium*, in the youngest subfamily Epidendroideae. Previous work on the expression of SI

in other families shows that pollen is generally inhibited in one of two places: either immediately, on the stigma itself, or within the style, preventing the pollen tube from growing fully to the ovules. These are *pre-fertilization* mechanisms, meaning that they occur before the pollen is ever able to reach the ovule. To determine if a plant is SI, one first needs to perform pollination experiments—manually pollinating the plant with its own pollen and manually pollinating the plant with pollen from another unrelated individual, documenting the resulting fruit set in both cases. Johansen (1990) did such experiments with an impressive 63 *Dendrobium* species, 44 of which would not set seed with their own pollen. He also noticed something unique to orchids: in these SI species, pollen was not inhibited on the stigma nor within the style. Most *Dendrobium* species take three months to set seed once pollinated, but in self-pollinated SI species, the flower is dropped from the plant (known as *abscission*) 3-21 days following self-pollination. Johansen argued that, while unusual, this mechanism still occurred pre-fertilization as the flower abscised long before the three months it would ordinarily take for the plant to produce seed. This argument is important, because *post-fertilization* events that occur after the pollen reaches the ovule are difficult to attribute to a particular "mechanism"—there are other reasons why a plant might lose an embryo that are not caused by SI. The most common reason being *inbreeding depression*.

A prevailing theory is that plants adapted to avoid self-fertilization because it is the most severe form of *inbreeding*, and inbreeding is bad for biological fitness. Of course, inbreeding can be bad: hemophilia was so pervasive in the largely inbred European monarchy during the 19th and 20th centuries that it was dubbed "the royal disease" (Rogaev et al., 2009). But why are genetic diseases more prevalent in inbred populations? Many deleterious mutations are *recessive*, meaning a diploid organism like us humans needs two copies of the gene—one from each parent—to suffer from the mutation's negative effects. Only one copy of the recessive mutant gene would be "hidden" (or *sheltered*) by the normally functioning healthy gene. Mating between unrelated individuals increases the odds that the recessive mutations present in either parent are not *the same* recessive mutations, and so their offspring will inherit at least one healthy copy of each gene. The problem for closely related individuals is that they have similar genomes and are more likely to carry copies of the same recessive mutations, meaning two mutant copies are more likely to reach their offspring. And if the closely-related individual you mate with is you—well then, you have the same genome as yourself, and you share all the same recessive mutations.

Populations that are accustomed to cross-fertilization have what is referred to as *sheltered load*: behind the "normal" copies of their genes lie many recessive mutations built up over generations. If an individual like this suddenly self-fertilizes, their offspring suffer the consequences of these hidden recessive mutations—for many cross-fertilizing plants, this might mean being unable to produce offspring from self-fertilization at all, because of so many lethal sheltered mutations. These plants are not really SI; there is no genetic "mechanism" underlying their inability to self-fertilize, just many, many generations of cross-fertilization. This is why the pre-fertilization distinction matters: if the rejection occurs post-fertilization, it's extremely difficult to distinguish SI from plain-old inbreeding depression. And, in many orchid species with especially complex mechanisms promoting cross-pollination—meaning, these species rarely self-fertilize—self-pollination does indeed result in greatly reduced seed set and seed viability (*Planthera leucophaea* examined in Wallace, 2003; *Pleurothallis* species in Borba et al., 2001). But let us suppose that a cross-fertilizing individual does produce at least a small amount of seed from self-fertilizing, and that these seeds then undergo several more generations of self-fertilization. Each time, offspring with two copies of negative recessive mutations may die early or otherwise fail to reproduce, slowly removing such mutations from the population entirely. This phenomenon is known as *purging* (**Figure 8**). Purging is natural selection at work, weeding out the deleterious recessive mutations until the population slowly becomes adapted to repeated generations of self-fertilization. And there are many immediate benefits available to plants that adapt to self-fertilization. No pollinators around? No problem! No suitable mates? No worries! Instead of only contributing half of your genes to your progeny, your progeny are now constructed entirely of your genes. But Darwin was stumped. He documented 23 orchids that exclusively pollinated themselves (known as *autogamy*), finding it enormously difficult to justify their existence given his claim that nature abhors perpetual self-fertilization. Autogamy in British *Ophrys apifera* was, in his own words, "perplexing in an unparalleled degree" (Darwin, 1862, pg. 71), and he argued that such a mode of reproduction must be uncommon in the family (**Figure 9**). Darwin missed the mark again. Catling (1990) reported 350 autogamous orchid species distributed throughout every subfamily, and we now know of at least 100 more.

Despite Darwin's confusion on this point, he was not entirely wrong. There is a problem with relying on self-fertilization in the long-term. (Long-term in the evolutionary sense, meaning thousands or millions of generations.) Decades of empirical work suggest that, in part, the answer lies in variation. While genetic mutation is the fundamental source of all variation, sex re-shuffles genes into many different combinations—increasing variation beyond what would be possible through mutation alone. This may be why sex has continued across the tree of life these last billion years. While self-fertilization



Figure 8: A cartoon depiction of purging. (a) A plant that ordinarily cross-fertilizes instead self-fertilizes. The plant produces few offspring, (b) many of which have mutations with negative fitness consequences reducing their chances for survival and successful mating. (c) Those that are able to survive and reproduce do, with the majority of seed coming from those plants with the highest biological fitness. (d) After several generations, the negative mutations are purged from the population leaving only those that are adapted to self-fertilization.



Figure 9: **Botanical illustration of autogamous** *Ophyrus apifera*, showing the (a) flower, (b) the column including the anthers and stigma, (c) the floral components with the petals removed, and (d) the pollinia and a cross-section of the ovules. The anther cells are positioned directly over the stigma, so that when the flower opens and pollinia are produced the plant can easily self-pollinate. (Digital illustration adapted from Thomé, 1885.)

is still sex, for every generation that an organism self-fertilizes, variation in its offspring is reduced by *half*. This can negatively affect the process of natural selection, which acts on existing variation—if there is no variation, there is no natural selection. Given enough generations of self-fertilization, all offspring would be more-or-less genetically identical. What happens when there is a sudden change in the environment? What happens when a new disease appears, ravaging the population? To respond to these changes quickly enough to avoid extinction, groups of organisms would need existing genetic variation—different genes and different combinations of these genes —or sufficient time, so that some members may adapt to changes through the process of natural selection. Unfortunately for orchids and all other living creatures, natural selection has no foresight: it cannot make predictions about the future; therefore, traits with benefits in the short-term like self-fertilization may still prevail.

Let's now return to our first question: what accounts for the incredible diversity of Orchidaceae? In truth, we don't know for sure. Since Darwin, many features of orchids have been proposed as drivers for the group's species richness: the evolution of pollinia packages, highly-specialized floral morphology particular to certain pollinators, the epiphytic growth habit, and predominant distribution in tropical regions, among others (Givnish et al., 2015). There is also evidence from other plant families that selfcompatible species tend to go extinct faster than those that obligately cross-fertilize (Goldberg et al., 2010), meaning that the various floral contrivances promoting cross-fertilization may aid in preventing the extinction of these species across large timescales. Importantly, the reproductive flexibility of the orchids is also suggested to account for the group's extraordinary diversity (Dressler, 1981). To test what traits may correlate with the diversity of certain groups, scientists examine the patterns of observed species distribution—that is, which related groups have more species and which have fewerand computationally evaluate whether or not there are features these species possess that may explain the observed patterns of their distribution. This is new and exciting research, only possible with increased computational power and large-scale datasets, practically unimaginable in Darwin's day.

Perhaps you might like to know the breeding systems of your favorite orchids, or you are wondering which species might produce viable seed from their own pollen and which will not. To start, I have compiled a publicly-available database from 356 published reports with information on 1107 unique orchid species and their breeding systems. This represents only four percent of the family, so many species are missing. I also recommend *Techniques for Pollination Biologists* (Kearns & Inouye, 1993), an excellent resource that outlines the various steps required to start evaluating the breeding system of a plant, allowing you to perform experiments on your orchids just as Darwin did for several decades. Darwin himself had only intended for his work on orchids to result in a single chapter of a book otherwise dedicated to examples of natural selection at work in both plants and animals. But, caught up in their extraordinary beauty, and more to the point, the unparalleled "beauty of the adaptations of parts" (Darwin, 1861a), *Orchids* ballooned to more than 400 pages, outlining mysteries that extend far beyond the orchids themselves, still grappled with by biologists today.

### 3.1 Cited Literature

- Aceto, S., & Gaudio, L. (2011). The MADS and the beauty: Genes involved in the development of orchid flowers. *Current Genomics*, *12*(5), 342–356.
- Alcock, J. (2006). An enthusiasm for orchids: Sex and deception in plant evolution. Oxford University Press.

- Allen, A., & Hiscock, S. (2008). Evolution and phylogeny of self-incompatibility systems in angiosperms. In *Self-incompatibility in Flowering Plants* (pp. 73–101). Springer.
- Borba, E. L., Semir, J., & Shepherd, G. J. (2001). Self-incompatibility, inbreeding depression and crossing potential in five Brazilian *Pleurothallis* (Orchidaceae) species. *Annals of Botany*, *88*(1), 89–99.
- Catling, P. (1990). Auto-pollination in the Orchidaceae. Orchid Biology. Review and Perspectives. V., 123–158.
- Correvon, H. (1899). *Cyprepedium calceolus*, L.: Yellow lady's slipper. In *Album des orchidées de l'Europe centrale et septentrionale*. Genève, Librairie Georg. **biodiversitylibrary.org/page/**15446363
- Darwin, C. (1859). On the Origin of Species: A fascimile of the first edition. Harvard University Press.
- Darwin, C. (1861a). Letter 3221: Darwin, C. R., To Hooker, J. D., 28 July Aug 1861. In F. Burkhardt (Ed.), *The Correspondence of Charles Darwin* (Two thousand, third, Vol. 9, p. 222). Cambridge University Press.
- Darwin, C. (1861b). Letter 3284: Darwin, C. R., To Darwin, W. E., 12 Oct 1861. In F. Burkhardt (Ed.), *The Correspondence of Charles Darwin* (Two thousand, third, Vol. 9, p. 302). Cambridge University Press.
- Darwin, C. (1862). On the various contrivances by which orchids are fertilised by insects, and on the good effects of intercrossing. John Murray.

Darwin, C. (1872). The effects of cross and self fertilisation in the vegetable kingdom. John Murray.

Dressler, R. L. (1981). The orchids: Natural history and classification. Harvard University Press.

- Givnish, T. J., Spalink, D., Ames, M., Lyon, S. P., Hunter, S. J., Zuluaga, A., Iles, W. J., Clements, M. A., Arroyo, M. T., Leebens-Mack, J., & others. (2015). Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proceedings of the Royal Society B: Biological Sciences*, 282(1814), 20151553.
- Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., & Igić, B. (2010). Species selection maintains self-incompatibility. *Science*, *330*(6003), 493–495.
- Harder, L. D., & Johnson, S. D. (2009). Darwin's beautiful contrivances: Evolutionary and functional evidence for floral adaptation. *New Phytologist*, *183*(3), 530–545.
- Igić, B., Lande, R., & Kohn, J. R. (2008). Loss of self-incompatibility and its evolutionary consequences. International Journal of Plant Sciences, 169(1), 93–104.
- Johansen, B. (1990). Incompatibility in *Dendrobium* (Orchidaceae). *Botanical Journal of the Linnean* Society, 103(2), 165–196.
- Kearns, C. A., & Inouye, D. W. (1993). Techniques for pollination biologists. University Press of Colorado.
- Larson, B. M., & Barrett, S. C. (2000). A comparative analysis of pollen limitation in flowering plants. Biological Journal of the Linnean Society, 69(4), 503–520.
- Linden, J. J. (1885a). *Catasetum discolor* Lind. In *Lindenia: Iconographie des orchidée* (Vol. 1). Gand: Impr. F. Meyer-van Loo. biodiversitylibrary.org/page/47502633
- Linden, J. J. (1885b). *Cattleya trianae* Lind. Var *memoria lindeni* L. Lind. In *Lindenia: Iconographie des orchidées* (Vol. 1). Gand: Impr. F. Meyer-van Loo. biodiversitylibrary.org/page/47502633
- Lloyd, D. G. (1965). Evolution of self-compatibility and racial differentiation in *Leavenworthia* (Cruciferae). *Contributions from the Gray Herbarium of Harvard University*, *195*, 3–134.

- Pérez-Escobar, O. A., Gottschling, M., Whitten, W. M., Salazar, G., & Gerlach, G. (2016). Sex and the Catasetinae (Darwin's favourite orchids). *Molecular Phylogenetics and Evolution*, 97, 1–10.
- Poinar, G., & Rasmussen, F. N. (2017). Orchids from the past, with a new species in Baltic amber. Botanical Journal of the Linnean Society, 183(3), 327–333.
- Quammen, D. (2007). The Reluctant Mr. Darwin: An intimate portrait of Charles Darwin and the making of his theory of evolution. WW Norton & Company.
- Rogaev, E. I., Grigorenko, A. P., Faskhutdinova, G., Kittler, E. L., & Moliaka, Y. K. (2009). Genotype analysis identifies the cause of the "royal disease". *Science*, *326*(5954), 817–817.
- Romero, G. A., & Nelson, C. E. (1986). Sexual dimorphism in *Catasetum* orchids: Forcible pollen emplacement and male flower competition. *Science*, *232*(4757), 1538–1540.
- Thomé, O. W. (1885). Ophrys apifera. In Flora von Deutschland Österreich und der Schweiz. biodiversitylibrary .org/page/12306734
- Tremblay, R. L., Ackerman, J. D., Zimmerman, J. K., & Calvo, R. N. (2005). Variation in sexual reproduction in orchids and its evolutionary consequences: A spasmodic journey to diversification. *Biological Journal of the Linnean Society*, *84*(1), 1–54.
- Wallace, L. E. (2003). The cost of inbreeding in *Platanthera leucophaea* (Orchidaceae). *American Journal of Botany*, 90(2), 235–242.
- Yam, T. W., Arditti, J., & Cameron, K. M. (2009). "The orchids have been a splendid sport"—an alternative look at Charles Darwin's contribution to orchid biology. *American Journal of Botany*, 96(12), 2128–2154.

Zimmerman, J. K. (1991). Ecological correlates of labile sex expression in the orchid *Catasetum viridiflavum*. *Ecology*, *72*(2), 597–608.

# **CHAPTER 4**

# SCIENTIFIC ADDENDUM TO THE ORCHIDS AND THEIR BREEDING SYSTEMS

#### 4.1 Abstract

One of the largest families of organisms on Earth, the orchids are known for their stunning diversity of plant-pollinator relationships. While most species in the family are thought to be self-compatible (SC) and thus able to set seed from their own pollen, self-incompatibility (SI)—the inability to self-fertilize—has been uncovered in several genera, most notably *Dendrobium*. Despite widespread reports of the SI phenotype in monocots, no common SI mechanism has been observed across the group as in eudicots—but several have been uncovered in other Asparagales families that are sister to Orchidaceae. There is an important need for breeding system data across every angiosperm family, especially monocots and in particular the orchids, a groups whose breeding systems are poorly studied and accounts for nearly eight percent of all flowering plant diversity. As breeding systems are factors known to influence macroevolutionary patterns, selection response, and trait distribution, it is especially vital to determine the patterns of such traits in this extraordinarily species-rich group of angiosperms.

To address this need, we compiled information on breeding systems in the orchid family from 356 published reports. Our current database contains 1310 entries, accounting for 1107 unique species and approximately four percent of the family. We find the same bimodal expression of SI in the family as uncovered previously in many others. The majority of SC species in the database are obligately self-pollinating, while most others are require a pollinator for successful fertilization. Most genera are

fully or mostly SC, with few genera studied displaying SI in all species. Determination of breeding systems poses many difficulties, but complications specific to orchids include the production of nonviable seeds following self-fertilization and extremely long generation times. We make this database publicly available so that orchid breeders, horticulturalists, or scientists may use this information to inform their work.

### 4.2 Materials and methods

Our primary database including breeding system and trait data was constructed with information from 356 published materials. Our phylogeny originated from the broadly inclusive seed plant phylogeny created by Smith & Brown (2018). After the various pieces of data were compiled, all binomials were standardized using the Taxonomic Name Resolution Service v5.0 (Boyle et al., 2021). For the analyses described below, the associated reference list can be found in **Appendix B**. All analyses were conducted in R (R Core Team, 2020).

### 4.2.1 Study system

Orchidaceae is the largest family of flowering plants, containing approximately 30,000 species not including cultivars and hybrids (Castillo-Pérez et al., 2019). The orchids have a cosmopolitan distribution and are found across many habitats, with the highest species-richness in the tropics. Tropical orchids are primarily epiphytes while terrestrial orchids grow in more northern latitudes. Recent fossil evidence suggests that the most recent common ancestor of extant orchids lived approximately 76–84 million years ago, diversifying rapidly following the mass extinctions at the K/T boundary (Ramírez et al., 2007). Molecular phylogenies indicate that Orchidaceae is sister to all other Asparagales families (Givnish et al., 2015). While highly-modified as individual species co-evolved with pollinators, orchid

flowers share similar features, including the fused column containing the male and female reproductive organs, the packaging of pollen in sugary sacs of pollinia, the modified labellum petal, and the rostellum (Aceto & Gaudio, 2011). Orchids commonly produce thousands of dust-like seeds that require fungi for proper germination.

## 4.2.2 Breeding system data

Using ISI Web of Knowledge Science Citation Index and Google Scholar, we collected published reports with dozens of search terms related to plant mating and breeding systems within Orchidaceae. We also included data from conservation reports, local flora and field guides, books, conference presentations, theses, and monographs. Breeding system data was collected and scored with a quality indicator (Grossenbacher et al., 2017). (See **Chapter 2** for more details on quality indicators.) Instances where relative success of selfed and outcrossed hand-pollinations were reported, or those that directly provided a measure along with transparent calculation procedures, received a score of one (highest quality). This high-quality data was used to calculate Index of Self-Incompatibility (ISI). ISI is defined as

$$ISI = 1 - \frac{relative \ selfed \ success}{relative \ outcrossed \ success}$$
,

where relative pollination success is defined as the ratio of fruits set to flowers pollinated (after Lloyd, 1968). Calculation of ISI with fruit set is ideal when fruit types allow, as reduced seed set can occur for reasons other than the prezygotic action of SI, notably inbreeding depression. While pollination experiments that reported relative fruit sets were preferred (298), when this was not available ISI was calculated with seed sets (n = 58). If only manual self-pollinations were performed but these crosses yielded fruit, these data were also assigned a quality of one (n = 110). Instances where hand pollina-
tions were performed but fruit set or seed set was not reported within the study were given a quality two, and those that simply stated the breeding system with no pollination experiments reported were given a quality of three. When reports stated the breeding system with an associated citation and that citation did not address the breeding system of that species, these data were not scored (i.e., "NA"). Our search included papers published through March 2019.

It is possible for ISI to take on negative values in the rare instances (n = 62) when the relative success of self-pollinations is higher than that of cross-pollinations. As outlined previously in Raduski et al. (2012), the existence of negative ISI values likely results from experimental sources of errors (low sample sizes, incompatibility of outcross pollen, mistakes). Within a single population, there is no clear biological basis for increased fruit set following self-pollination versus cross-pollination (Lloyd & Schoen, 1992). Consequently, we set all negative values of ISI to zero. Species with ISI values above 0.8 are generally classified as SI (Bawa, 1974).

#### 4.2.3 Other trait data

Concurrently with breeding system data for individual species, we collected information on other traits if such traits were also reported in the study. Growth habit (epiphytic, terrestrial, lithophytic, etc.), geographic location, and altitude were almost always included within the original published reports, but in the rare instances they were not, this information was collected from another published report when available. Data were also collected on the following when available: pollination syndromes, specialized floral morphology and nectaries, viable seed set from self and outcross pollen, and pollen limitation calculated from observations of natural and manual pollination success.

#### 4.2.4 Phylogenetic relationships

Our Orchidaceae tree was extracted from the individual clade data used to construct the broadly inclusive seed plant phylogeny from Smith & Brown (2018). This phylogeny of the order Asparagales was constructed using 6972 taxa in NCBI (retrieved from http://www-personal.umich.edu/ ~eebsmith/big\_seed\_plant\_datasets/trees/ on July 27, 2019). We used a list of all recognized Orchidaceae genera from TPL to quickly prune for relevant taxa. Species designations were then formally synonymized using the procedure below.

We created two trees. The first was a species-level tree containing only those species also included in our breeding system database. We also used the same Asparagales phylogeny to construct a genus-level Orchidaceae tree by randomly selecting and retaining one tip from each genus. Both trees were made ultrametric by using penalized likelihood with the chronopl function in the R package ape (Paradis et al., 2004).

## 4.2.5 Taxonomic name resolution

We synonymized species names for all binomials in breeding system (n = 1107) and phylogenetic datasets (n = 4604) using the Taxonomic Name Resolution Service v5.0 (Boyle et al., 2021) with three sources selected: USDA (National Plant Data Team, 2020), TPL (The Plant List, 2013), and Tropicos (Missouri Botanical Garden, 2020). Our search settings allowed partial matches, and a match accuracy that returned all names found within an edit distance of four (four missed letters, four transposed letters). For both datasets, species designations classified as "Accepted" were kept. Species designations classified as "Synonym", "Illegitimate", "Invalid", "No opinion", "Misapplied", or "Rejected name" were replaced with "Accepted" names when available or removed from the datasets. We also removed species no longer recognized in the family Orchidaceae.

## 4.3 Results

Our final dataset contains data from 356 published reports across six languages including books, theses, monographs, conservation reports, flora and field guides, conference presentations, and empirical papers. Our collection includes work from 1862 through 2019. There are 1310 entries containing data for 1107 unique species across 295 genera. Of these, 1210 entries include some information on species' breeding systems, and 951 uniquely characterize species as SI or SC. Our dataset contains 477 entries of quality one data, 131 of quality two data, and 603 of quality three data.

TABLE III: **Summary of SI distribution by orchid subfamily.** For each subfamily: the number of unique species for which data on breeding system is available (Species), the percent of incompatible species of those tested in each subfamily (Percent SI), and the percent of the subfamily tested (Percent coverage).

SUBFAMILY	Species	Percent SI	Percent coverage
Apostasioideae	1	0	6.7
Vanilloideae	33	3	18.3
Cypripedioideae	17	6.3	10.3
Orchidoideae	252	3.2	6.7
Epidendroideae	648	26.5	3
Total	951	19.1	3.8

We find that slightly fewer than 20 percent of tested Orchidaceae species are SI and that the highest proportion of SI species appear in Epidendroideae (Table III). We report 3.6 percent coverage in Epidendroideae, 7.3 percent coverage in Orchidoideae, 18.9 percent coverage in Vanilloideae, 13.3 percent coverage in Cypripedioideae, and 6.7 percent coverage in Apostasioideae with a total family-wide coverage of 3.6 percent. We were unable to find reported data for 350 genera across the 545 genera included on our phylogeny. The vast majority of genera with reported breeding system data have such data for less than 20 percent of the genus (**Figure 10**). Of the 194 genera included in both our tree and dataset, 162 genera are either all SI or all SC (**Figure 11**).

## 4.3.1 Index of self-incompatibility

Our literature search of experimental crossing data yielded a measure of ISI for 467 species (Figure 12). ISI fruit was preferred (n = 305), but ISI calculated from seed set was used where fruit yield was not reported (n = 62). This data spans all subfamilies and 84 genera. We also added species where only manual self-pollinations were performed but full fruit set was observed, assigning an ISI value of 0 (n = 110). Thus, our total ISI dataset included 467 entries from 137 genera (18 percent of 763 recognized genera). As observed across angiosperms (Grossenbacher et al., 2017; Raduski et al., 2012), the frequency appears strongly bimodal, with 82 percent of species possessing values of ISI  $\leq 0.2$  or ISI  $\geq 0.8$ . For species where ISI calculation was possible, approximately one-fifth (20.8 percent) are classified as SI using a cutoff value of 0.8.

## 4.3.2 Phylogenetic distribution of and frequency of SI

There is a low frequency of SI across every subfamily except the youngest and largest, Epidendroideae (Figure 13). Many reports originate from a single genus, *Dendrobium*, where 82 species have been tested with 52 reported as SI. However, reports of SI in Epidendroideae span an additional 57



Figure 10: **Genus-level orchid phylogeny displaying breeding system data coverage.** Each circle is scaled to represent the relative size of the genus within the family. Color gradient within circles represents 0-100 percent breeding system data coverage within that genus. Genera were manually binned into each of six percent-coverage categories:  $0 (n = 350), 0.01-2 (n = 38), 2.01-4.5 (n = 36), 4.51-12 (n = 43), 12.01-30 (n = 43), and 30.01-100 (n = 34). The largest category included 350 genera with 0 percent coverage. Our molecular phylogeny places Apostasioideae as sister to all other subfamilies, then Cypripedioideae, with Vanilloideae sister to Orchidoideae and Epidendroideae. While we retain these relationships here, more recent molecular phylogenies show strong support for reversing the relationship between Cypripedioideae and Vanilloideae, with Cypripedioideae being sister to the younger two subfamilies (Givnish et al., 2015). Thus we label the node of all three subfamilies. In all molecular phylogenies, Orchidoideae and Epidendroideae share a common ancestor and are widely accepted be the youngest subfamilies with a stem age of approximately 64 My. The tree originates from the broadly inclusive seed plant phylogeny of Smith <math>\vartheta$  Brown (2018).



Figure 11: **Proportion of each orchid genus that is SI or SC.** Numbers in parentheses indicate how many species from each genus were tested. Of the 194 genera represented here, 162 are all SC (146) or all SI (16).



Figure 12: **Distribution of ISI values in Orchidaceae from collected literature (number of species = 467).** ISI measures the strength of SI reaction by comparison of relative selfed and outcrossed success of fruit set per flower after manual pollination. An ISI value below 0.8 is classified as self-compatible, and above 0.8 is classified as self-incompatible. Species with negative ISI values were set to zero (n = 62).

genera across the group, with 70 percent of SI species coming from, in addition to *Dendrobium*, *Callista* (n = 22), *Restrepia* (n = 21), *Oncidium* (n = 13), and *Cattleya* (n = 12).

To the best of our knowledge, only seven studies have examined expression of SI following selfpollination (Table IV). Most originate from a single study, Johansen (1990), where 44 *Dendrobium* species aborted selfed ovules within 3-21 days following self-pollination—the rejection occurring far earlier than the three months it ordinarily takes these species to set seed. Niu et al. (2017) examined *Dendrobium longicornu* and *D. chrysanthum*—the only two species with reported pollen tube arrest within the style—for the presence of molecular markers characteristic of the widespread SI system found across many core eudicot families, S-RNase-based GSI. These markers were not present. Other common reports include irregular pollen tube growth or slower growth of selfed pollen tubes than outcrossed pollen tubes. No work on the genetic basis of SI in orchids currently exists, likely due to the production of thousands of dust-like seeds and extremely long generation times.

### 4.3.3 Other breeding systems

Most SC orchids are either herkogamous (n = 173), often with extremely elaborate pollination mechanisms that are highly-specific to certain pollinators, or entirely self-pollinating (n = 421). Autogamous orchids are widespread across every subfamily, while herkogamous orchids are found in every subfamily save Apostasiodeae. Other SC species exhibit unique pollination mechanisms: observations include self-pollination triggered by rain (Fan et al., 2012; Suetsugu, 2019) or even changes in floral morphology facilitating self-pollination after a certain period of herkogamy (Liu et al., 2006). There are nine species across Epidendroideae (n = 5), Orchidoideae (n = 3) and Vanilloideae (n = 1) that produced some seed asexually in the absence of pollen, but agamaspermy generally seems rare within



Figure 13: **Species-level orchid phylogeny with breeding system data (number of species = 388).** Species found on both the tree and in the breeding system dataset, where black represents SI and gray represents SC. Those species with conflicting calls (SI and SC, with low-quality indicator of 3) were not included. Species from the subfamilies Vanilloideae, Cypripedoideae, Orchidoideae, and Epidendroideae are all represented: only one report of SI is found across Vanilloideae, Cypripedoideae, and Orchidoideae. The vast majority of SI reports are found scattered across the largest and youngest subfamily, Epidendroideae. The tree originates from the broadly inclusive seed plant phylogeny of Smith & Brown (2018).

		0.1
Species	Site of Inhibition	Study
Octomeria crassifolia	Stigma; irregular pollen tube growth	Barbosa et al., 2009
Octomeria grandiflora	Stigma; irregular pollen tube growth	Barbosa et al., 2009
Octomeria praestans	Stigma; irregular pollen tube growth	Barbosa et al., 2009
Anathallis heterophylla	Stigma; irregular pollen tube growth	Gontijo et al., 2010
Anathallis microphyta	Stigma; irregular pollen tube growth	Gontijo et al., 2010
Anathallis rubens	Slowed growth of selfed tubes in style	Gontijo et al., 2010
Anathallis sclerophylla	Stigma; irregular pollen tube growth	Gontijo et al., 2010
Pleurothallis johanensis	Weak SI; irregular pollen tube growth	Borba et al., 2001
Pleurothallis teres	Weak SI; irregular pollen tube growth	Borba et al., 2001
Pleurothallis ochreata	Weak SI; irregular pollen tube growth	Borba et al., 2001
Pleurothallis adamantinensis	Strict SI; irregular pollen tube growth	Borba et al., 2001
Pleurothallis fabiobarrosii	Strict SI; irregular pollen tube growth	Borba et al., 2001
Acianthera proilfera	Ovule abortion	de Melo et al., 2011
Dendrobium aciculare	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium acinaciforme	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium aloefolium	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium alterum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium aphyllum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium bicameratum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium blumei	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium brevimentum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium cariniferum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium chrysotoxum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium crumenatum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium denudans	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium devoniannum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium disticum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium draconis	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium ellipsophyllum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium falconeri	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium farmeri	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium gratiostissumum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium griffithianum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium hendersonii	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium indivisum var. pallidum	Ovule abortion; flower abscission	Johansen, 1990
Dendrobium keithii	Ovule abortion; flower abscission	Jonansen, 1990
Dendrobium lamellatum	Ovule abortion; flower abscission	Jonansen, 1990
Denarobium leonis	Ovule abortion; flower abscission	Jonansen, 1990
Denarobium linaleyi	Ovule abortion, flower abscission	Johansen, 1990
Denarobium linguella	Ovule abortion, flower abscission	Johansen, 1990
Denarobium mannii	Ovule abortion, flower abscission	Johansen, 1990
Denurobium moschatum	Ovule abortion; nower abscission	Johansen, 1990
Denurobium mucronatum	Ovule abortion; nower abscission	Johansen, 1990
Dendrobium pachyglossum	Ovule abortion; nower abscission	Johansen, 1990
Dendrohium pachyphyllum	Ovule abortion: flower abscission	Johansen 1000
Dendrobium pandurijerum	Ovule abortion: flower abscission	Johansen 1000
Dendrobium parcum	Ovule abortion: flower abscission	Johansen 1000
Dendrobium planihulhe	Ovule abortion: flower abscission	Johansen 1000
Dendrobium padagravia	Ovule abortion: flower abscission	Johansen 1000
Dendrohium primulinum	Ovule abortion: flower abscission	Johansen, 1990
Dendrobium primuinum	Ovule abortion: flower abscission	Johansen 1000
Dendrobium sacundum	Ovule abortion: flower abscission	Johansen 1000
Dendrobium setifolium	Ovule abortion: flower abscission	Johansen 1990
Dendrohium subulatum	Ovule abortion: flower abscission	Johansen 1000
Dendrohium thursiflorum	Ovule abortion: flower abscission	Johansen 1000
Dendrobium virgineum	Ovule abortion: flower abscission	Johansen 1990
Dendrohium chrysanthum	Style	Niu et al 2017
Dendrobium longicornu	Style	Niu et al. 2017
Restrenja brachvnus	Stigma: irregular pollen tube growth	Milner et al 2015

TABLE IV: **Site of fertilization failure following self-pollination for the studied orchid species.** To the best of our knowledge, seven studies have examined the location of pollen tube arrest following self-pollination in self-incompatible species. Of the 61 species represented here, 46 are from the genus *Dendrobium*. In SI *Dendrobium* species, floral abscission is the most common response, while irregular pollen tube growth is found across other groups.

104

orchids. Most reports of sexual dimorphism or dichogamy come from a single genus, *Catasetum*, where reports of moneocy, dioecy and dichogamy have all been observed. Often, environmental factors determine sexual expression in these plants (Pérez-Escobar et al., 2016).

#### 4.4 Cited Literature

- Aceto, S., & Gaudio, L. (2011). The MADS and the beauty: Genes involved in the development of orchid flowers. *Current Genomics*, *12*(5), 342–356.
- Bawa, K. S. (1974). Breeding systems of tree species of a lowland tropical community. *Evolution*, *28*(1), 85–92.
- Boyle, B. L., Matasci, N., Mozzherin, D., Rees, T., Barbosa, G. C., Kumar Sajja, R., & Enquist, B. J. (2021). Taxonomic name resolution service, version 5.0. *Botanical Information and Ecology Network*. https://tnrs.biendata.org/
- Castillo-Pérez, L. J., Martínez-Soto, D., Maldonado-Miranda, J. J., Alonso-Castro, A. J., & Carranza-Álvarez, C. (2019). The endemic orchids of Mexico: S review. *Biologia*, *74*(1), 1–13.
- Chase, M. W., Cameron, K. M., Freudenstein, J. V., Pridgeon, A. M., Salazar, G., Van den Berg, C., & Schuiteman, A. (2015). An updated classification of Orchidaceae. *Botanical Journal of the Linnean Society*, *177*(2), 151–174.
- Fan, X.-L., Barrett, S. C., Lin, H., Chen, L.-L., Zhou, X., & Gao, J.-Y. (2012). Rain pollination provides reproductive assurance in a deceptive orchid. *Annals of Botany*, *110*(5), 953–958.
- Givnish, T. J., Spalink, D., Ames, M., Lyon, S. P., Hunter, S. J., Zuluaga, A., Iles, W. J., Clements, M. A., Arroyo, M. T., Leebens-Mack, J., & others. (2015). Orchid phylogenomics and multiple drivers of

their extraordinary diversification. *Proceedings of the Royal Society B: Biological Sciences, 282*(1814), 20151553.

- Grossenbacher, D. L., Brandvain, Y., Auld, J. R., Burd, M., Cheptou, P.-O., Conner, J. K., Grant, A. G., Hovick, S. M., Pannell, J. R., Pauw, A., & others. (2017). Self-compatibility is over-represented on islands. *New Phytologist*, *215*(1), 469–478.
- Johansen, B. (1990). Incompatibility in *Dendrobium* (Orchidaceae). *Botanical Journal of the Linnean Society*, 103(2), 165–196.
- Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, *34*(7), 1812–1819.
- Liu, K.-W., Liu, Z.-J., Huang, L., Li, L.-Q., Chen, L.-J., & Tang, G.-D. (2006). Self-fertilization strategy in an orchid. *Nature*, *441*(7096), 945–946.
- Lloyd, D. (1968). Pollen tube growth and seed set in self-incompatible and self-compatible *Leavenworthia* (Cruciferae) populations. *New Phytologist*, *67*(1), 179–195.
- Lloyd, D. G. (1965). Evolution of self-compatibility and racial differentiation in *Leavenworthia* (Cruciferae). In R. Rollins & R. Foster (Eds.), *Contributions from the Gray Herbarium of Harvard University* (Vol. 195, pp. 3–134). The Grey Herbarium of Harvard University.
- Lloyd, D. G., & Schoen, D. J. (1992). Self-and cross-fertilization in plants. I. Functional dimensions. International Journal of Plant Sciences, 153(3, Part 1), 358–369.

Missouri Botanical Garden. (2020). Tropicos. http://www.tropicos.org

National Plant Data Team. (2020). USDA, NRCS: The PLANTS Database. http://plants.usda.gov

- Niu, S.-C., Huang, J., Xu, Q., Li, P.-X., Yang, H.-J., Zhang, Y.-Q., Zhang, G.-Q., Chen, L.-J., Niu, Y.-X., Luo, Y.-B., & others. (2018). Morphological type identification of self-incompatibility in *Dendrobium* and its phylogenetic evolution pattern. *International Journal of Molecular Sciences*, *19*(9), 2595.
- Niu, S.-C., Huang, J., Zhang, Y.-Q., Li, P.-X., Zhang, G.-Q., Xu, Q., Chen, L.-J., Wang, J.-Y., Luo, Y.-B., & Liu, Z.-J. (2017). Lack of S-RNase-based gametophytic self-incompatibility in orchids suggests that this system evolved after the monocot-eudicot split. *Frontiers in Plant Science*, *8*, 1106.
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*(2), 289–290.
- Pérez-Escobar, O. A., Gottschling, M., Whitten, W. M., Salazar, G., & Gerlach, G. (2016). Sex and the Catasetinae (Darwin's favourite orchids). *Molecular Phylogenetics and Evolution*, *97*, 1–10.
- Raduski, A. R., Haney, E. B., & Igić, B. (2012). The expression of self-incompatibility in angiosperms is bimodal. *Evolution*, 66(4), 1275–1283.
- Ramírez, S. R., Gravendeel, B., Singer, R. B., Marshall, C. R., & Pierce, N. E. (2007). Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature*, *448*(7157), 1042–1045.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
- Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. American Journal of Botany, 105(3), 302–314.
- Suetsugu, K. (2019). Rain-triggered self-pollination in *Liparis kumokiri*, an orchid that blooms during the rainy season. *Ecology*, *100*(7), e02683.

The Plant List. (2013). The Plant List (Nos. v1.1). https://www.theplantlist.org/

# INTERLUDE

So far this thesis has been concerned with the examination of a single adaptation in flowering plants that has profound consequences for the group's evolution. It is the kind of complexity that is mostly appreciated by scientists. But the work we do as scientists eventually becomes folded into our university pedagogy. How do we appropriately prepare students to appreciate the kind of complexity that is fundamental to the study of evolutionary biology? In the next two chapters, I wrestle with this question by examining the concept of adaptation as it relates to the development of student understanding.

# CHAPTER 5

# UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION. I. A FRAMEWORK FOR SYSTEMATIC ANALYSIS.

## 5.1 Abstract

An enormous body of work demonstrates that students struggle to understand evolutionary theory in general and natural selection in particular. Here we aim to uncover the basic structure of student explanations in the content domain of adaptation, in the broadest possible terms. We ascribe to recent work in the philosophy of biology that suggests appropriate explanations must cite causal mechanisms along with relevant information on the patterns such processes produce. Thus we start by developing a framework that encompasses and integrates two important and understudied features necessary for reasoning about adaptation: (1) distinguishing between patterns and processes, and (2) distinguishing between patterns and processes at both the individual level and the species level. We use the resultant framework to systematically assess student thinking in a study that spanned two semesters and all levels of biological study. We find that students struggle to both formulate robust explanations and to integrate information across levels. We suggest that these results stem from a diversity of causes, including underlying cognitive processes and the general pedagogy of biological sciences. Students need more support in the classroom to develop their own coherent and robust explanatory frameworks.

#### 5.2 Introduction

Evolution is the light that illuminates the science of biology (Dobzhansky, 1973). Yet for decades, a large body of work has accumulated demonstrating that students hold tenacious misconceptions regarding evolutionary principles and the theory of evolution by natural selection (Ferrari & Chi, 1998; Greene Jr, 1990; Gregory, 2009; Nehm & Schonfeld, 2008). Rather than simply cataloging misconceptions, more recent work aims to uncover the complexities of reasoning in these areas and the underlying cognitive processes that may account for difficulties in understanding evolution (e.g., Coley & Tanner, 2015; Ha et al., 2019; Trommler et al., 2018). Open-ended assessments like the Assessment of Contextual Reasoning about Natural Selection (ACORNS) instrument (Nehm et al., 2012) provide insights into the structure of conceptual understanding by examining how students construct explanations for biological phenomena. Building a coherent framework that allows for deep conceptual understanding of adaptations—the heritable traits resulting from the process of natural selection—requires that students integrate lines of seemingly counter-intuitive reasoning across many levels of biological organization.

Mayr (1961) famously notes that the historical nature of biological sciences means that there are appropriate causal explanations at two broad levels: the level of the individual (proximate) and the level of the species (ultimate). Consider two causal explanations for the existence of woody tissue in plants: (1) secondary growth caused by cell division in the vascular cambium (proximate), and (2) inter- and intra-specific competition for light over time (ultimate). Of course, these processes produce observable patterns. Woody tissue exists both because there are trees, and because trees, like all organisms, must adapt to their environment to survive. Thus there are proximate and ultimate causes, but also corresponding proximate and ultimate *patterns*. The distinction between the proximate and ultimate is essential to understanding the action of natural selection, but so too is the distinction between patterns and the causal processes responsible for such patterns.

We know that Charles Darwin formulated his theory of evolution by natural selection without ever understanding (or perhaps, misunderstanding) the mechanisms of inheritance. He did so by focusing on the adaptations (traits) themselves and their distribution within environments: he was one of the first to recognize a process that may yield the observed pattern of organisms that appear engineered for specific tasks, interfacing with their surrounding ecology using remarkably honed sets of tools (Vincent & Brown, 2005). This pattern Darwin observed may be thought of as a description of *what is*, an essential step in determining *why it is so*. But what is the *why*? Here, the *why* is an explanation that provides a causal mechanistic justification. For instance, that a tree is made of wood is an observable pattern, or a description of *what is*. Explaining the existence of wood in *that* tree requires citing the mechanism of cell division in the vascular cambium. This is a paired pattern and processes that produce wood.) But the existence of wood in a single tree does not explain the existence of wood across vascular plants. This species-level pattern—that wood has evolved repeatedly in many groups—has a specieslevel causal explanation. The grid below outlines and organizes the distinctions between descriptions and explanations, and their various levels, into a simple framework (**Figure 14**).

Explaining biological phenomena is clearly no easy task, especially for students in the early years of formal biological study. We know that students hold tenacious misconceptions related to natural selection but the basic structure of student reasoning is still relatively unclear. While a large body

	Pattern (static)	Process (dynamic)	
	Material description	Material explanation	
Proximate (individual-level)	Description of what the trait is or made of, or the organism that possesses it	Explanation of the individual-level processes responsible for the trait	
Ultimate (species-level)	<i>Evolutionary description</i> * Description of the function or adaptive significance of the trait for the species	<i>Evolutionary explanation</i> Explanation of the historical evolutionary processes responsible for the trait	

Figure 14: **A framework for organizing biological descriptions and explanations.** Patterns are descriptions of *what is* that stem from our observations. Processes are mechanisms that provide causal explanations—explain why—the pattern exists. In biology, both patterns and processes can occur broadly at the level of the individual and at the level of the species. \*An evolutionary description—identifying the adaptive value of a trait—is essential in discovering the evolutionary processes that may account for the trait. However, evolutionary descriptions are distinct from explanations and themselves do not have explanatory power. When provided in the place of a causal explanation, these descriptions become teleological; as descriptions they are merely teleonomic. See text for more details on these distinctions.

of work examines student construction of causal or mechanistic explanations (e.g., Southerland et al., 2001; Abrams & Southerland, 2001; Chi et al., 2012; Russ et al., 2009; Scott et al., 2018; Speth et al., 2014), very few empirical studies examine the which students provide descriptions or explanations in response to questions that allow for answers at varying levels of biological organization. To the best of our knowledge, only one previous study examines this feature of student responses. Riemeier (2009)

interviewed ten secondary school students with questions aimed at causal explanation and found that 59% of responses were instead evolutionary descriptions (descriptions of adaptive value, termed "functional descriptions" by the author) rather than causal explanations, and that the most frequently generated explanations were teleological.

But an interesting feature of teleological reasoning about biological traits is that it comes in two broad flavors: using the current function of a trait to justify the trait's existence (not legitimate as causal explanation), and using the current function of the trait as a clue about the possible selective pressures influencing the development of the trait over time (legitimate as pattern identification). The second can become a legitimate scientific explanation if it invokes the causal process of natural selection. Kampourakis (2020) refers to the the first kind of teleological explanation as "design teleology," in which explanations stem from the perception of design in organisms, while the latter he refers to as "selective teleology." Similarly, Trommler & Hammann (2020) distinguish between these two facets of teleological reasoning as "ontological teleology" and "epistemological teleology" — in this case, the first describes the *cause* of a trait being its functionality (not legitimate), while the latter simply identifies a trait in terms of its effect (legitimate).

For this distinction we use the traditional terms of "teleological" and "teleonomic." Originally coined in 1958, "teleonomic" is meant to emphasize the fact that a description of an end-directed mechanism does not imply that the end result is the cause of the mechanism itself (Pittendrigh, 1958). This word is useful in biological systems, where natural selection favors those traits that enhance survival and reproduction of individuals with them relative to those without. In a sense, examining the "end result" of such a process provides valuable clues as to the action of natural selection over time. But it

is important that the end result is not conflated as a cause in and of itself. An important distinction is that a teleonomic description, when used in the place of a causal explanation, can become teleological. As Mayr (1991) notes, there is no evidence for the idea that natural selection itself is a goal-directed, although it does result in phenotypes that were "selected for" over time. Thus the word emphasizes the idea that the end result of a trait is important information for inductive reasoning, but it need not imply causality for the trait itself. A trait emerging from a teleonomic process can be seen as serving a function such as a fit of form and function, but it emerges from trial and error not by forethought or goal-directed design (Crespi, 2000; Schofield, 2021).

Indeed, evolutionary or teleonomic descriptions of traits (as opposed to evolutionary explanations) are essentially "ultimate patterns" with enormous utility in the study of evolution, allowing evolutionary biologists to examine what historical processes may be responsible for them. For example, documenting the current distribution of traits across groups of organisms can provide insight into the kinds of selective pressures responsible for the maintenance of such traits across macroevolutionary timescales (e.g., Goldberg et al., 2010). Take for instance the following pair of explanations (from Trommler et al., 2018), marked by the authors as teleological and causal:

Sweating occurs because ...

- 1. ... in this way, the body can eliminate excess heat (*teleological*).
- 2. ... the muscles that surround the sweat glands contract (*causal*).

The first, as a causal *explanation*, is not appropriate. However, as a statement of teleonomic *description*, it is indeed true that the body eliminates excess heat via perspiration. Moreover, the observed patterns of evaporative heat-loss mechanisms in mammals—and the enormous variation in these

mechanisms—provides valuable fodder for examining the evolutionary mechanisms influencing thermoregulation in mammalian organisms (Robertshaw, 1971). That is, the observation that the body eliminates excess heat through sweating is not only correct, but an important step in understanding the trait itself and its possible adaptive significance. Recent work argues that a robust understanding of evolution requires that students are explicitly aware of the distinctions between teleonomic and teleological reasoning (Hammann & Nehm, 2020). The distinction between pattern (description) and process (explanation) is also essential for students to recognize in forming a robust understanding of evolution. Such skills are necessary for students to appreciate the large body of work in the field of evolutionary biology that wrestles with patterns produced via the process of natural selection. Without the ability to distinguish between the facets of pattern and process at different levels, students will lack crucial skills necessary for making sense of the study of evolution at large.

For biologists, a formal distinction outlining the various levels of explanation required to fully understand a trait first appeared in the form of Tinbergen's Four Questions (Tinbergen, 1963). He described four possible levels of explanation (similar to Aristotle's Four Causes): developmental and mechanistic explanations relating to individual organisms, and phylogenetic and adaptive explanations relating to groups of organisms or species. These questions were organized into a grid by Nesse (2013), providing enormous utility as an explanatory framework by addressing both proximate and ultimate causality, along with corresponding proximate and ultimate patterns necessary for making inferences about the associated processes responsible for generating them. We borrow heavily from the resultant grid, which is valuable in that it organizes the levels of biological description and explanation into a very clear framework that can be used as a tool for systematically assessing student thinking as it relates to adaptation. Moreover, by clearly delineating the possible levels of description and explanation, the framework leaves room for appropriate biological description that takes adaptive value into account (i.e., ultimate pattern descriptions or teleonomic reasoning). We can then use this framework to develop a broad view of the various descriptions and explanations, at the level of the individual and at the level of the species, that account for a given adaptation (**Figure 15**).



Figure 15: **Using the framework to examine the adaptation of woody tissue.** Here we demonstrate how the adaptation of woody tissue in plants may outlined using the proposed framework. We can see patterns (or descriptions) at both the individual and species level, and the corresponding processes (or explanations) that provide causal mechanisms accounting for the existence of such patterns. Of course, there are likely many more processes occurring at both levels that account for the observed patterns of woody tissue, but the framework provides a broad view of the type of information that must be integrated and a method for integrating it.

Students demonstrate their understanding of concepts by providing explanations. What is the basic structure of such explanations in the content domain of adaptation? Here, we aim to determine the extent to which students provide genuine explanations (causal mechanisms) versus descriptions (identifying patterns) related to adaptation, and whether such responses more commonly address the proximate (individual level) or the ultimate (species level). We use this framework as a means to systematically assess student responses for this purpose. We distinguish between responses that correctly identify the teleonomic ultimate pattern (evolutionary description) versus those that use illegitimate teleological reasoning (e.g., the internal motivation of the organism). We describe the use of this framework—along with preliminary validity and reliability evidence—at a large Midwestern research university in a study that spanned two semesters and several levels of biological study.

#### 5.2.1 Research questions

We first begin the process of establishing validity and reliability evidence for the framework and the associated questionnaire. These two components are interrelated but differ in important ways. Establishing the validity of the framework requires gathering evidence that the framework accurately conceptualizes the manner in which we plan to categorize student responses. Establishing the validity of the questionnaire is more complex, and is likely not possible in the scope of a single study. We begin here by focusing on two components of validity evidence for the questionnaire—content and substantive validity evidence, described below. Our first series of research questions relate to both validity and reliability evidence gathered in this study for both components.

1. Is the framework conceptualized accurately? (1.1) Does the questionnaire sufficiently represent the content domain of adaptive traits at the species level (content validity evidence)? (1.2.1)

Does the questionnaire produce the intended reasoning patterns in students (substantive validity evidence)? (1.2.2) Is there agreement in how student answers are placed within the framework (reliability evidence)? (1.3)

We then administer the questionnaire and use the associated framework to categorize student responses. From those results we address the questions below.

- 2. Do students more commonly invoke descriptions (provide patterns) or explanations (provide causal mechanisms) when asked causal "why?" questions? (2.1) What level do their answers more commonly address—individual-level patterns and processes (proximate) or species-level patterns and processes (ultimate)? (2.2)
- 3. What proportion of answers invoke teleological versus telonomic reasoning?
- 4. Does previous exposure to formal biology instruction influence the propensity of students to provide descriptions or explanations at the proximate or ultimate level (3.1) or invoke teleological reasoning (3.2)?

## 5.3 Theoretical background

## 5.3.1 Conceptual framework: Measurement

There exist many methods for establishing the validity and reliability of educational instruments. Here we align with the *Standards for Educational and Psychological Testing* (AERA et al., 2014), where reliability describes the extent to which measurement is replicable, consistent, and error-free, and validity describes the extent to which instrument interpretations are supported by evidence and relevant theory (as in Sbeglia & Nehm, 2020). Our questionnaire is not an assessment per se, in the sense that (1) there are many correct answers for each question, and (2) the questionnaire does not produce a "score." The purpose is to assess students' biological thinking as it relates to adaptation, not to evaluate correctness. Questions are designed such that there are possible answers at both the proximate and ultimate level. While the questions ask for causal justifications, we do not categorize pattern recognition (providing descriptions) as "incorrect" as students are rarely taught the features of robust and causal explanations (see Chapter 6 for more details). Unlike other types of educational instruments, we do not suggest that certain responses are indicative of latent constructs, i.e., we are not using the questionnaire to measure adjacent features like conflict or acceptance.

Thus to assess validity of our questionnaire, we focus on categories of evidence that align best with our measurement purpose. The construct validity framework we use includes six categories of evidence, including (1) evidence based on test content (that it aligns with the domain it is intended to represent), (2) evidence based on response processes (that it engages appropriate cognitive response), (3) evidence based on internal structure (relationship between test items and their conformity to the construct being measured), (4) evidence based on relationships to other variables (if the test claims to measure some specific variable, evidence that it does so), (5) validity generalization (can the test be generalized to new situations), and (6) evidence of consequences (evidence on the soundness of proposed interpretations for the test). We focus here on the first two categories of validity evidence: evidence based on test content (i.e., content validity), and evidence based on response processes (i.e., substantive validity).

To establish content validity evidence, we performed two main tasks: a literature review and semistructured interviews. We first conducted a literature review in the content domain of natural selection and adaptation, specifically related to epistemological and pedagogical aspects in this area. We also examined previous studies that empirically assessed student understanding of natural selection and adaptation. We used this literature review and expert judgement as a guide to develop our questionnaire. We developed a series of open-ended questions that allowed students to provide responses without being limited by a pre-determined set of choices. We constructed such questions in a manner that allowed for responses at both the individual and species level. While we purposely constructed the questions with "why?" language, (i.e., casual questions), their open-ended nature allowed students to provide either descriptions or explanations as responses. We then conducted semi-structured interviews with a small (n = 10) group of students using this set of open-ended questions.

To establish substantive validity evidence, we performed a pilot version of the study in the Summer Semester prior to finalizing our questionnaire. We carefully evaluated student responses and modified questions that the majority of students misinterpreted. We used student responses as a guide to formalizing our coding framework: student responses largely addressed various patterns and processes at both levels. We also performed external expert validation, where a panel of three biologists were asked to interpret the questions and provide responses. We used expert responses as a tool to determine if our questions appropriately addressed the reasoning patterns such questions were designed to elicit. We also asked that the framework be evaluated to determine if it appropriately conceptualized the construct we intended to measure (the various descriptions and explanations appropriate for adaptive traits).

To establish reliability evidence, we produced a random sample of twenty responses that were provided to two biology educators unfamiliar with this study or the questionnaire. The educators were provided the framework, along with a rubric for placing responses in their appropriate categories (i.e., coding). The coding results from the two educators was compared to the original coding to determine agreement.

## 5.3.2 Conceptual framework: Epistemological and pedagogical aspects

## 5.3.2.1 The concept of adaptation

The concept of adaptation is tightly linked and often used interchangeably with that of natural selection in educational materials, so it is worth taking some time to tease apart and define these terms. Simply put, natural selection is a process whereby heritable traits that increase an organism's ability to survive or reproduce will increase in frequency in the population. The traits themselves are adaptations. But we may also dive more deeply into this concept to appreciate some of its inherent complexity. Natural selection is a process that occurs in groups that contain replicating units, where units can be anything from genes to whole populations. In this context, four necessary conditions reveal when the process of natural selection will occur (reproduced from Stearns, 2013):

- 1. Replicating units that possess some feature must vary in the amount of replicants they produce.
- 2. The feature itself must vary between units within the same population.
- 3. There must be a nonzero correlation between the feature and the successful production of replicants.
- 4. The feature must be heritable.

Given these conditions, features that are positively correlated with the production of replicants will increase in frequency over time. Now, what is a feature? A feature is a character or trait that is distinct and observable in an organism or across a group of organisms (e.g., morphological, physiological, or

behavioral aspects, gene sequences, or emergent properties like geographic ranges). Such a feature is adaptive when the unit that possesses it replicates and survives more, thus increasing in frequency at the expense of others that do not possess the feature.

There are two important but separate aspects related to the concept of adaptation: the first is the business of defining an adaptation and the second is the business of demonstrating that a specific trait is indeed adaptive. Both are tricky, but the second is likely more controversial and outside the scope of this paper. We begin with defining the concept of adaptation. Biologists tend to agree that an adaptation is a product of the process of natural selection, the "feature" possessed by organisms that causes increased frequency of offspring over generations. It is not always clear whether or not the concept of adaptation requires history. For example, a historical definition of adaptation would dictate that a feature is adaptive because natural selection has acted upon it in the past. An ahistorical definition only requires that the feature provide a reproductive advantage to an organism in its current environmental context. Of course, both can be true, but they need not always be. These broad definitions clearly emphasize different aspects of adaptation that may overlap or not depending on the kind of explanations being sought.

The real utility in definitions and concepts is in their application to practical problems. In biology especially, we use concepts like natural selection and adaptation to propose questions about the mechanisms that shape the natural world. But definitions and concepts vary in their utility. Take for instance the plethora of species concepts proposed over the last century (Schemske, 2000): we know that the biological species concept (BSC) does not account for all aspects of "species" as we observe them on Earth. But the great utility of the BSC is that its definition proposes a clear research program, because reproductive isolation is something that can be evaluated empirically. Thus, historical definitions of adaptation pose challenges because it is not always possible to infer historical processes based on current patterns. Nevertheless, phylogenetic comparative methods applied with increased computational power and large datasets allow biologists to formulate explanations regarding historical processes that were previously unanswerable (e.g., Beaulieu & O'Meara, 2016; Maddison et al., 2007; Mayrose et al., 2011). The important takeaway is that both broad definitions of adaptation, when applied, result in questions of a different character that require different methods of evaluation. This results in varied explanations that may or may not overlap.

The purpose here is not to argue in favor of one definition over the other, but to demonstrate some of the murky features of biological concepts and how such nuances may impact student understanding of adaptation. A cursory review of biology textbooks indicates that there is little consistency in how adaptation is defined and explained (Kampourakis, 2013). The primary differences in definition stem from whether or not the definition includes historical processes or not. But the unique structure of our framework means that we need not pick one definition at the exclusion of the other, making it possible to uncover student thinking as it relates to historical versus ahistorical definitions. A feature may provide a selective advantage in the current environment (evolutionary description), and/or it may be the result of the historical process of natural selection (evolutionary explanation). Determining how students conceptualize adaptations generally allows us to address the nuances of such issues in the classroom.

#### 5.3.2.2 Biological explanation

Philosophers of science devote serious attention to the nature of explanation in science and across various disciplines. Here we summarize some of the epistemological aspects of explanation in general and biology explanation in particular. We describe and expound on some of the peculiar features of biological sciences that make codified rules for explanation particularly challenging. Finally, we discuss more recent attempts to unify explanation in biology under the umbrella of causality.

The philosophical debate regarding the nature of explanation in science was largely non-existent until the years following the publication of Hempel & Oppenheim (1948). Prior to this, the realm of science was largely concerned with testing and confirming hypotheses in the pursuit of uncovering general laws that make prediction possible. Put another way, science observed phenomena and worked to answer how such phenomena arose. But the dominant public perception was that science did not explain *why* such phenomena arose: explaining why was considered to be the realm of metaphysics or theology, outside the scope of scientific inquiry. With the formulation of the deductive-nomological (D-N) model, Hempel & Oppenheim (1948) demonstrated that it is indeed possible—as many scientists had done previously—for science to offer "true" explanations. In the decades that followed, a vast body of literature has accumulated in pursuit of a model of scientific explanation (see Salmon, 2006 for review). Three broad issues characterize the continued work in this area today (Braillard & Malaterre, 2015):

- What are the unique features of scientific explanations and can such features be developed into necessary and sufficient conditions?
- 2. Is causation a necessary component of explanation, and what model of causation is appropriate?

#### 3. What is the role of "context" in explanation?

Several models of explanation have been put forth since the D-N model (e.g., the inductive-statistical model, the statistical-relevance model, the causal-mechanical model), but a common characteristic of all models are constitutional flaws emanating from various features of the models themselves. Which is to say that none of the models fully capture all aspects of scientific explanation. This fact has resulted in two broad views toward the future: a unified theory of explanation that may stem from common aspects of previous models or entirely new ones (unification), and pluralistic approach with different models that may be discipline- or question-dependent (pluralism).

The pluralistic approach is prevalent in the field of biology, where the kinds of questions asked along with their explanations—differ in character from those of the physical sciences and thus do not fit neatly into previously-developed models. Unlike the physical sciences, there are far fewer general laws that can be cited to explain phenomena. Moreover, questions related specifically to the process of natural selection and adaptive traits may probe at entirely different features of these concepts and do so using entirely different methods. For instance, there are questions relating to the current adaptive value of a trait in its environment (ahistorical approach), but also the historical processes that led to the development of that trait over time (historical approach). Some questions may ask about the genetic changes that need to occur for the trait to appear, while others examine abiotic or biotic selective pressures stemming from the surrounding environment. Biologists may use methods that involve genetic sequencing, computational modeling, field observations, or greenhouse experiments. Perhaps one is interested in interspecific floral color polymorphism in sister species and the role natural selection plays in its maintenance. Table V summarizes a small sample of the types of questions that could be

QUESTION	Approaches
What genes are responsible for flower color in these species?	Transcriptome analysis (Sánchez-Cabrera et al., 2021)
How common are mutations and in what genes?	Comparative survey of known mutations (Streisfeld & Rausher, 2009)
What is the biochemical pathway responsible for flower color?	Biochemical analysis (Forkmann & Ruhnau, 1987)
How does a mutation alter this pathway?	Gene sequencing; cloning (Habu et al., 1998)
What are the evolutionary and ecological consequences of changes in floral color?	Field work; phylogenetic comparative analyses (Durbin et al., 2003)
Do pollinator shifts result in reproductive isolation (RI) between species?	Field work; greenhouse experiments (Ramsey et al., 2003)
What is the selective advantage of one flower color versus another?	Phylogenetic comparative analyses; greenhouse experiments (Bradshaw & Schemske, 2003)

TABLE V: **Questions and approaches on floral color polymorphism.** Adaptive traits are the result of complex processes and biological tradeoffs that interact at many levels of biological organization. Understanding an adaptation involves posing many questions in different areas of biological sciences that involve disparate and varying approaches. For example, floral color shifts have important ecological and evolutionary consequences in the observed diversity of angiosperms. Here we list just a small sample of questions related to this feature and the various approaches that may be employed to answer such questions.

asked and the possible approaches that could be employed to answer them. The fundamental takeaway is that the data that provides explanatory power—what characterizes a "good" explanation—depends heavily on the question for which an answer is sought.

The kinds of questions one can ask in biology leads to a plurality of approaches employed to answer them. It would not be appropriate to use field experiments to uncover the biochemical pathways responsible for flower color, nor would molecular approaches shed light on the ecological consequences of changes in such traits. Put another way, adaptive traits emerge from complex processes and biological tradeoffs at many levels, from alterations in nucleotide sequences to intra-specific competition. (Ignoring momentarily the challenges inherent in *proving* a trait is adaptive.) Whole subdisciplines are structured in practice around entirely different features of natural selection and resultant adaptations. To complicate matters, "ecology" and "evolutionary biology" are often treated as separate subdisciplines with unique questions and approaches. This is a curious state of affairs, given that evolution occurs only in environmental contexts and that the relationship between organisms and their environment is necessarily the result of evolutionary processes. There is no evolution without environment and no biotic environment without evolution. What does this mean for explanatory models in the domain of natural selection?

More recently, philosophers of biology have put forth models that center around mechanistic or causal accounts of explanation. While there is some debate among philosophers of science, there is general agreement that causality is an important—if not central—aspect of scientific explanation (Potochnik, 2013). Thus appropriate explanation would center around citing causes for specific features to be explained. Mechanistic accounts stem from the observation that many biological explanations appeal to mechanisms as opposed to general laws. These kinds of explanations are attractive because they contain relevant causes along with organizational information. For instance, to explain the existence of proteins one may cite the relevant mechanisms of DNA transcription and translation (Bechtel & Abrahamsen, 2005). To provide a mechanistic explanation, one must specify each component and its operation within the system—its temporal and organizational sequence—that linearly results in the feature to be explained (Bechtel, 2011). Clearly, mechanistic explanations have more utility in some areas of biology, notably cell biology, where relevant phenomena are often the result of observable and distinguishable mechanisms that occur in a sequential manner. However, mechanistic explanations are challenging when the feature to be explained is the result of many mechanisms that do not necessarily operate sequentially. This is especially true if we are interested in explaining the origin and maintenance of transcription and translation across the tree of life. In these cases, causal explanations may be preferred.

However, providing a causal explanation for adaptation becomes increasingly complex as there are a great deal of causes and/or mechanisms that may result in one. The wing of a penguin, for instance, is shaped by (1) biological trade-offs between flight and diving efficiency, (2) environmental pressures, and (3) developmental constraints. Focusing on any one of these areas may yield varying explanations. Mutations occur that provide the raw material for phenotypic change, but it is the reproductive success of particular organisms—shaped by access to resources and mates—that ultimately explains the persistence or development of an adaptation. Providing explanation based on developmental processes versus ecological processes will certainly involve citing different causes. Moreover, even if I am able to provide all the causes responsible for penguins' wings, it is unlikely that the same feature would emerge again if a similar bird species migrated to the same environment. That is, the explanation for penguins' wings is not necessarily generalizable to similar organisms in the same circumstances. Given these complexities, some philosophers of biology have put forth an account of explanation that cites causal *patterns* along with the particular causal influence(s) as a means of providing relevant scope to the explanation itself (Potochnik, 2020). Causal patterns provide the relevant conditions for which the particular explanation holds true. This kind of explanation holds power in the realm of adaptation where processes at many levels—highly dependent on context—are responsible for the existence of an adaptive trait. Moreover, this flavor of explanation fits neatly within the context of our framework.

#### 5.3.2.3 Biological explanation in biology education

Equipping students with the skills to develop explanations is a major goal of science education. Despite this, it seems there is no clear consensus on how to establish such a skill. Studies that examine student explanations do so with varying frameworks that do not always rely on work in philosophy of science. Responses are coded with some consistency across studies but also include important differences in the handling of teleological reasoning. Types of questions and the appropriate explanations for such questions are often not clearly defined. Here we summarize some of the previous work done in this area and highlight some of these issues.

Science is about more than the rote accumulation of facts and instead aims to provide insight into the world around us. Descriptive knowledge has important differences from explanatory knowledge. Thus equipping students with the skills to formulate scientific explanations is a central goal of science education. Given the lack of clear consensus on what constitutes an explanation, work in this area approaches the task of assessing student explanations in various ways. Some explicitly borrow from philosophy of science to formulate frameworks (e.g., Russ et al., 2009; Kampourakis & Zogza, 2008), while others use frameworks from other areas of scientific inquiry (e.g., Speth et al., 2014 uses SBF systems theory). Despite the varying frameworks, nearly all studies examining students' biological explanations argue that such explanations must be causal, and generally mechanistic. Which is to say, a causal explanation is one that cites a relevant biological mechanism. No studies integrate causal mechanisms with associated patterns that provide scope for the explanation being produced.

We know that many mechanisms may account for a biological feature, thus the scope of causal dependence is handled with slight differences across studies. However, varying studies drawing on disparate frameworks still maintain some commonalities: mechanisms can be cited at different levels (e.g., proximate/ultimate or genetic/phenotypic), and usually teleological explanations are included as a separate category (Kampourakis & Zogza, 2008; Southerland et al., 2001; Speth et al., 2014). However, Abrams & Southerland (2001) noted that when asked "how?" questions (probing for causal or mecha-

nistic answers), students often redirected with a "why?" answer (addressing a rationale that could be used to explain). From the interview snippets provided, these "why?" answers are appear to generally take the form of evolutionary description of the phenomena in question (e.g., "(the plant) need(s) light to grow, so it grows that way", pg 1275). This differs from other studies that would characterize this kind of response as teleological.

Recent work highlights the fact that teleological reasoning differs from teleonomic reasoning in important ways that have not previously been appreciated. What makes this especially challenging is that evolutionary descriptions—teleonomic reasoning—can also occur at multiple levels. Thus much of the previous work that has categorized student explanation as "teleological" may be applying this term incorrectly to evolutionary descriptions at the population or species level. It may be that some portion of teleological explanations are actually teleonomic descriptions that are not only correct, but provide relevant scope for addressing the trait in question. Yet there are very few empirical studies that examine students' ability to distinguish relevant patterns (i.e., provide descriptions) and provide corresponding explanations (i.e., describe processes or mechanisms). Studies primarily focus on whether or not students cite processes or mechanisms and in general categorize all evolutionary descriptions as teleological.

A major issue appears to be a lack of clarity in the kind of answer being sought by educators. A "whyquestion" seeks a relevant cause—and most educators and philosophers tend to agree this means citing an appropriate mechanism or process. In contrast, a "what-question" probes for descriptive answer. But if an important aspect of citing relevant causes is providing scope for the phenomenon in question, then associated descriptions are also necessary. How often are these distinctions made explicit for stu-
dents? There is no clear answer in the literature. While may studies note the difficulties children have in identifying "plausible physical causes" (Abrams & Southerland, 2001) and that teleological or anthropocentric reasoning is common (Bishop & Anderson, 1990; Coley & Tanner, 2015; Demastes et al., 1995; Tamir & Zohar, 1991), it is not clear when or how children are actually taught to identify questiontypes and appropriate explanations to address them. Given the plurality of question-types common in the discipline of biology—and the various methods used to address such questions—it is unlikely students will learn to appropriately address questions asked of them without explicit intervention from educators, but there is limited work in this area.

# 5.3.2.4 Conceptual understanding of natural selection

Assessing conceptual understanding of natural selection is challenging for three broad reasons (Anderson et al., 2002). First, there are several necessary components of natural selection that must be appropriately integrated. CANS, the Conceptual Assessment of Natural Selection (Kalinowski et al., 2016), outlines five broad categories: evolutionary change (including trait loss and gain, role of environment, and role of individual change), mutation, inheritance, selection (including environmental stress and the role of chance), and variation. These individual concepts exist within a network: which is to say, the concepts are all related and interdependent. Assessing conceptual understanding thus involves identifying students' understanding of each component along with their ability to reason across components. Of course, all these components are indeed necessary for understanding natural selection, but what is missing is students' ability to distinguish (1) which of these features is a pattern versus a causal process, and how those concepts are related, and (2) which of these components occur strictly among individuals, at the species level, or both. We know of no empirical study that examines students' ability to distinguish between these features in an integrated manner. This is an important omission, because the patterns produced by natural selection exhibit striking differences at the population level, dependent on whether selection is directional, stabilizing or disruptive.

Second, students' conceptions are rarely composed entirely of scientific (normative) or non-scientific (non-normative) conceptions: in fact, they are generally heterogeneous and composed of *both* normative and non-normative ideas (Nehm & Reilly, 2007; Nehm & Schonfeld, 2008). While certain types of reasoning may be common, particular combinations of normative and non-normative ideas are unlikely to be shared between individual students. The questions themselves may expose varying degrees of understanding as item features and context play an important role in the kind of reasoning invoked by students. Nehm & Ha (2011) identifies five specific challenge areas: (1) immunity or resistance versus other kinds of trait changes, (2) within- versus between-species differences, (3) trait loss versus trait gain, (4) animals versus plants, and (5) familiar versus unfamiliar traits or taxa. Thus generalizing knowledge about evolution across both varying contexts and the different nested concepts can be especially challenging.

Third, the concept of adaptation is not treated uniformly within courses. In educational materials, an adaptation may be considered (1) a process, (2) a trait resulting from the historical process of natural selection, or (3) a trait that provides a selective advantage in the current environment regardless of historical processes. A review of ten textbooks by Kampourakis (2013) shows a variety of definitions within the text, largely ahistorical, some process-based, and rarely mentioning natural selection. In general, the focus is on function: while an important aspect of adaptation, this exacerbates the tendency to reason about traits in teleological ways. Teleological explanations are some of the first used and understood by young children (e.g., *Why did you go to the store? Because we needed to buy bread.*) so students may be prepared early in school to provide these kinds of explanations in general (Lennox & Kampourakis, 2013). Given that textbooks—and likely classroom activities—are emphasizing evolutionary (functional) descriptions of adaptation without the relevant action of natural selection, it may be quite challenging for students to move away from teleological tendencies.

Given these points, assessing conceptual understanding of adaptation remains a difficult exercise. In order to design instructional materials that best interact with students' prior knowledge, it is important to understand common features of student responses. On the one hand, previous work indicates that student responses are largely heterogeneous and approach problems in varied ways, which would suggest that finding such commonalities may not be possible. However, empirical studies that examine student responses to open-ended questions with a broad and integrated framework—encompassing patterns and processes at two levels—are largely absent. Much of the complexity related to explanation and the concept of adaptation reviewed above points to something we biologists have known for centuries: uncovering the historical action of natural selection and diagnosing adaptation is a complicated pursuit, largely stemming from the complex interplay between processes occurring at many levels of biological hierarchy in varying contexts. To help our students develop the cognitive skills necessary to make sense of adaptation requires that we first see how they make sense of it now—in the broadest possible terms, so that our instruction impacts the greatest number of students. We suggest that the proposed framework offers the ability to gain such an understanding.

#### 5.4 Materials

### 5.4.1 Participant sample

Participants were drawn from two semesters (Fall 2020 and Spring 2021) across four biology courses at a large, public, research-oriented university in the Midwestern United States (n = 584 for the preand post-assignment [36.5 percent participation rate]). See Table VI for a summary of the four courses and their general content focus: courses varied in their focus on evolution, although all courses use evolution as a core idea. Academic information collected from registrar data included undergraduate class standing (freshman, sophomore, junior, senior), major (binned into biology-related majors [biological sciences, bioengineering, biochemistry] and other majors), and prior biology coursework. When available, we gathered information on where participants graduated from high school for a broad overview of student demographics. Participant information is summarized in Table VII, and distribution of previously earned credits is shown in Figure 16. The sample of participants included students with varying experience in biological sciences, majors and non-majors, from diverse regions in the Chicagoland area.

#### 5.4.2 Framework development

We first developed a series of 15 open-ended "why?" questions related to adaptation at the species level, informed by a literature review and expert judgement. We administered these questions in a pilot study during the Summer 2020 semester (n = 33). We used these questions as a heuristic to evaluate the general categories of responses from students to aid in selecting an appropriate framework. We observed student responses included descriptions as well as explanations, and that they commonly addressed either proximate or ultimate levels. This fit comfortably within the organized Tinbergen (1963) framework developed by Nesse (2013) (with slight modifications). In addition, we found that

	Students	Description
Evolution		Advanced undergraduate. Theories of evolution and heredity, major patterns
Fall	13	and trends in evolution, speciation, biodiversity, and evolutionary medicine,
Spring	17	with an introduction to data science.
Animal Physiological Systems		Intermediate undergraduate. Basic physiology of animals in the context of the
Fall	124	themes of homeostasis and evolution, focusing on how organisms adjust and
Spring	93	respond to changes in their environments.
General Ecology Laboratory		Intermediate undergraduate. Application of ecological and evolutionary con-
Fall	38	cepts with hands-on experiments and field trips to local natural areas.
Spring	61	
Biology of Populations and Communities		Introductory undergraduate. Scientific skills, evolution, Mendelian and popu-
Fall	136	lation genetics, biological diversity, and ecological systems including ecosys-
Spring	102	tem processes and human impacts.

TABLE VI: **Description of each course and student number by semester.** Student participants were drawn from four courses at all levels of biological study. Here we report course descriptions and the number of participants from each course by semester.

	Fall	Spring
<b>Biological Sciences</b>		
Freshmen	13	18
Sophomores	50	45
Juniors	76	64
Seniors	74	38
Other		
Freshmen	20	38
Sophomores	44	40
Juniors	17	18
Seniors	17	12
Totals	311	273

TABLE VII: **Breakdown of grade level by major and semester.** We bin all biological sciences related majors (biological sciences, bioengineering, biochemistry) into one group. All other majors represented under **Other**.



Figure 16: **Distribution of the number of previously earned credits by student participant (number of students = 579.** The total number of previously-earned credits in biological sciences from both withininstitution and transfer were added for each student. Credit earnings range from 0-43; mean number of credits = 11.2; mode = 0.

the framework we developed also accounted for relevant aspects addressed in previous studies, including historical, ahistorical, process-based, and teleological explanations versus teleonomic descriptions of adaptation.

#### 5.4.3 Questionnaire development

The original 15 open-ended questions were developed to address adaptation at the species level. For most questions, we specifically used the word "why?"—this is in contrast to other studies which avoid the use of this phrasing because it is shown to often yield teleological responses. However, if providing causal explanations is a central goal of science education then it is essential that students can appropriately interpret and respond to "why" questions. Our questions were developed with the item features identified in Nehm Ha (2011) as a guide: we included questions about both plants and animals, trait loss and trait gain, unfamiliar taxa, and within- and between-species variation. Our questions included four broad content areas of adaptation at the species level:

- 1. Convergent evolution (one example of convergent trait loss and one example of convergent trait gain)
- 2. Developmental constraints (in mammals and in birds)
- 3. Macroevolutionary patterns of diversity (species diversity and emergent features like species ranges)
- 4. Generation and maintenance of variation (within and between species)

After the pilot study, we altered questions that the majority of students misinterpreted. Following the Fall 2020 semester, we found certain questions still yielded confused answers so we altered these questions (e.g., a question about bird coat color polymorphism in male birds was repeatedly interpreted as a sexual selection question). Following the administration of the questionnaire in the Fall and Spring, we conducted semi-structured interviews with a group of students (n = 10) to follow-up on their understanding of the questions. See **Appendix C** for the full questionnaire, including questions that were altered between the Fall and Spring Semesters, or only included in the Spring Semester.

## 5.4.4 Administration

The questionnaire was made available to students during the first two weeks of classes as an extra credit assignment on Blackboard (learning management software platform). Each student was shown a random sample of five questions from the larger pool of ten. Because the goal of the questionnaire

was to assess thinking patterns and not correctness, students were free to return to the assignment as many times as they liked during this period, and were encouraged to use and cite resources to support their claims. At the end of the two week period, the assignment was closed. Students were then asked to repeat the assignment once more during the last two weeks of classes with the same instructions. They again saw a random sample of five questions from the larger pool of five. We expected that item order had no impact on student responses as the questions were unrelated to each other.

#### 5.5 Methods

We administered the assignment in both semesters at the beginning and end of the course. We downloaded the data from Blackboard into a spreadsheet and assigned random numeric identifiers for individual students. We coded each response based on the framework and noted teleological answers. Such answers were either (1) goal-directed, in that organisms evolved for a particular purpose, or (2) based on the internal motivation of the organism itself. Students that only completed the pre-assignment were not included: we constrained results to only those students that completed both. We performed all subsequent data visualization and analysis in R (R Core Team, 2020).

## 5.5.1 RQ1: Validity and reliability

**Content validity evidence.** We conducted a literature review that encompassed the concept of adaptation and conceptual understanding of natural selection, along with the epistemological nature of explanation in science and in biology generally. We also assessed studies that explicitly examined student construction of biological explanation. We assembled a panel of students for semi-structured interviews where students were shown a random sample of four questions from the pool of ten. Our student sample included two freshmen, seven sophomores, one junior; five of the students were in biology-related majors and five that were not. We asked students to (1) read aloud the question, (2) explain what they believed the question was asking, and (3) provide an answer to the question. We then asked follow-up questions to assess how alterations in the wording of the question may yield better alignment between what the question attempted to elicit and how students understood it. The average interview time was 20m57sec.

**Substantive validity evidence.** We used the results from our literature review and carefully evaluated student responses from the summer pilot study. We edited or removed questions that were misinterpreted by the majority of students based on the justification(s) they provided in their answers. We also sent our framework and associated questionnaire to three biologists and asked them to provide answers to a random sample of four questions. We used this information to establish whether or not our questions targeted the appropriate reasoning patterns.

**Reliability evidence.** Two biology educators were given a random sample of 20 answers along with a rubric for coding questions. The interrater reliability was calculated along with Fleiss' kappa (Fleiss, 1971) from the R package irr (Gamer et al., 2012).

# 5.5.2 RQ2: Level of descriptions and explanations

Results were coded using the associated framework into the categories outlined in Table VIII. We evaluated the distribution of responses before and after instruction in the Fall and Spring semester by comparing these distributions using two-sample Kolmogorov-Smirnov tests. We focused on the following comparisons:

Category	DESCRIPTION	Example
Proximate patterns Individual-level static descriptions	Description of what the trait is made of, or the or- ganism that possesses it.	Flowers have different colors because they contain different levels of pigments like carotenoids.
Proximate processes Individual-level dynamic explanations	Explanation of the individual-level processes re- sponsible for the trait	Flowers have different colors because of coor- dinated differential expression of the carotenoid biosynthetic genes.
Proximate processes & Proximate patterns Individual-level dynamic explanations	Combination of a proximate description and proxi- mate explanation used to justify response.	Flowers have different colors because they have dif- ferent genes and proteins. When they mate with dif- ferent individuals, they produce offspring that also have different genes and proteins.
Proximate processes & Ultimate patterns Species-level descriptions & individual-level explanations	Combination of a proximate explanation and ulti- mate description used to justify response.	Flowers have different colors because of coordi- nated differential expression of certain genes, allow- ing them to attract different pollinators.
Proximate patterns & Ultimate patterns Individual-level descriptions & individual-level explanations	Combination of a proximate description and ulti- mate description used to justify response.	Flowers have different colors because they have dif- ferent genetic makeup and proteins. There are also flowers in different environments that attract differ- ent pollinators.
Ultimate pattern Species-level static descriptions	Description of the function or adaptive significance of the trait for the species.	Flowers have different colors because they attract different pollinators.
Ultimate processes Species-level dynamic explanations	Explanation of the evolutionary history that ac- counts for the trait.	Flowers have different colors because of pollinator- mediated selection on floral color over time.

TABLE VIII: **Categories for response coding with descriptions and examples.** Responses were coded into one of seven categories. Descriptions of each category and example responses are listed here.

- The overall distribution for response one across both semesters: compared to the distribution of response one in the Fall Semester, the distribution of response one in the Spring Semester, and the overall distribution of response two.
- 2. The overall distribution of response one in the Fall Semester: compared to the distribution of response one in the Spring Semester, and the distribution of response two in the Fall Semester.
- 3. The overall distribution of response one in the Spring Semester: compared to the distribution of response two in the Spring Semester.
- 4. The overall distribution of response two in the Fall Semester: compared with the overall distribution of response two in the Spring Semester.

We used the overall distribution of responses to determine if descriptions or explanations were more commonly paired with proximate or ultimate level responses. We visualized results with the R package vcdExtra (Friendly, 2021).

#### 5.5.3 RQ3: Teleological versus teleonomic justifications

While evaluating student answers, teleological justifications were noted in tandem. Teleological justifications differ from teleonomic ones in that they are more than simply functional descriptions of traits, but make statements that invoke the motives or aims of organism, or the needs of organisms, as a justification for the existence of a trait. See the most commonly observed teleological justifications coded in Table IX. We also marked other kinds of non-normative reasoning in a separate category (e.g., confusion regarding inheritance, timescales, common ancestry, or plants). Thus we had three categories of reasoning: scientific, teleological, or other, where other included all non-normative justifications that

Teleological Reasoning Pattern	Example Response
Personal motivations or aims of the organism	Penguins needed to swim so their adapted their wings.
Needs enforce changes on organism directly (Lamarckian in flavor)	Cave fish did not need their eyes so they got rid of them.
Observed pattern of diversity on Earth is predetermined or static	There are so many different species on Earth because every species has a job; lions can't pollinate flowers.
Natural selection is goal-oriented	There are so many different species on Earth because if there were only one it would go extinct.

TABLE IX: **Common teleological reasoning patterns and examples of student responses.** Teleological reasoning patterns present in student responses largely fell into one of these four categories. Example responses from each category are also included.

were not teleological. We also compared the distribution of teleological responses using two-sample Kolmogorov-Smirnov tests as outlined above in RQ2.

We used the overall distribution of responses to determine what kind of reasoning (teleological, scientific, or other) was more commonly paired with descriptions or explanations. We visualized results with the R package vcdExtra (Friendly, 2021).

## 5.5.4 RQ4: Effects of previous exposure

We gathered additional data from the registrar on students' major and previously earned credits in biological sciences. We categorized various majors as biology-related or other. We combined all previously-earned credits in biological sciences at the postsecondary level from both transfer credits and within-institution credits. To visualize the distribution of descriptions and explanations across credit levels, we binned these into six levels: 0 credits, 1-4 credits, 5-11 credits, 11-20 credits, 21-30 credits, and 31+ credits.

We compared the overall distribution of response one pair-wise to the distributions observed when the data was split up (1) by major, (2) by enrolled course, or (3) by number of previously earned credits, and (4) by question using two-sample Kolmogorov-Smirnov tests. We used the same procedure to examine the distributions observed within response two. We then compared the distributions observed between response one and response two for each of the variables.

We used logistic regression models (glm function in R with binomial distribution) for categorical variables to assess whether students' responses as a dependent variable correlated with their major, course, number of previously earned credits (binned), or assignment question as explanatory variables. We performed separate analyses for the Fall and Spring Semesters. We also used logistic regression models to assess possible correlations between teleological reasoning (dependent variable) and major, course, number of previously earned credits (binned), or assignment question. In both cases, we AIC model selection to distinguish among a set of possible models using the R package AICcmodavg (Mazerolle & Mazerolle, 2017).

# 5.6 Results

#### 5.6.1 RQ1: Validity and reliability

**Substantive validity evidence.** We used 15 open-ended questions in the summer pilot study and carefully evaluated student responses. We found that six of these questions invoked inconsistent responses from students that suggested fundamental misunderstanding of what the question intended to elicit. In the remaining nine questions, we found that each one included student responses that fell into our coding framework and that included scientific reasoning. We found that our expert panel produced responses to questions that agreed with our intent and our coding rubric.

**Content validity evidence.** Our literature review established that while natural selection and adaptation have long been examined by the education research community, studies that do so with an integrated framework that take descriptions and explanations at proximate and ultimate levels into account are largely absent. We used previous work to design questions that include known difficulties, like reasoning about trait loss versus trait gain and the sensitivity of students to item surface features. We used work in philosophy of science to construct questions that request causal mechanisms as justification. We also worked to create questions with possible responses at both the individual and the species level. In our semi-structured interviews, we find that all students correctly identify questions as relating to the topic of adaptation and provide general overviews of what the question is asking that fit in line with the intended construct. We still find that students struggle to provide causal explanations and instead prefer pattern-based descriptions. With more follow-up questions, students showed increased ability to integrate knowledge of proximate processes and ultimate patterns in their discussion of the question.

**Reliability evidence.** We report and overall interrater reliability of 83 percent with a Fleiss kappa of 0.456 (p < 0.01), indicating moderate agreement.

# 5.6.2 RQ2: Level of descriptions and explanations

Our final dataset includes 579 students that completed both the pre- and post-assignment, for a total of 2166 responses in both the pre- and post-assignment (4332 total responses). Five students that completed the assignment in multiple classes were removed. Using a threshold of p < 0.25 in two-sample Kolmogorov-Smirnov tests, we find no significant statistical difference in the distributions across any of our comparisons (between response one and two, semesters, majors, credit level, or course). Thus, we collapse response one and two from the Fall and Spring semesters (**Figure 17**). One additional question was added in the Spring Semester and wording was altered in another three questions. In addition, there are slight (but not statistically significant) differences in the distributions

between questions, so we compared these distributions separately (**Figure 18**). We find that the majority of students invoke ultimate descriptions in their responses, but in certain questions proximate processes are the most common (Q8 in response one; Q5, Q6, and Q8 in response two). We also observe that descriptions are overwhelmingly at the ultimate level while explanations are overwhelmingly at the proximate level in both responses (**Figure 19**).

#### 5.6.3 RQ3: Teleological versus teleonomic justifications

We find that while specific questions invoke statistically varied distributions of teleological justification, the distribution of teleological justifications does not differ between semester, majors, credit level, or course (see **Figure 17** and **Figure 18** for proportions of teleological justification). Justifications were not uniform for particular questions between response one and two (**Figure 20**, p < 0.01). Questions 1 and 3 invoked consistently more teleological justifications, while Questions 9 and 10 varied by response. We observe that a far greater proportion of descriptions are teleological, while a greater proportion of explanations are scientific (**Figure 21**, p < 0.01).

#### 5.6.4 RQ4: Effects of previous exposure

Using a cut-off of p < 0.25, the distribution of responses does not differ statistically between major, course, or number of previously earned credits in biology within or between responses. In our Spring Semester logistic regression, type of response (e.g., ultimate pattern) was best correlated with the question as an explanatory variable in both response one and response two. In the Fall Semester, the best performing models for response one included major (47 percent; AIC = 415.42), question (22 percent; AIC = 416.94), and major plus credit level (14 percent; AIC = 417.86) as explanatory variables. In response two, type of response was best correlated with the question only. Thus, with the





exception of response one in the Fall Semester, the question was the single best predictor of response type. Reasoning (teleological, scientific, or other) was also best explained by question as an explanatory variable across all semesters and responses.



Figure 18: **Overall distribution of student responses by question in Spring and Fall semesters (number of responses = 2166; number of students = 579).** Proportion of students that invoked each category of response by question are displayed here. Bars are filled in to represent proportion of those categories that invoked teleological reasoning. While this displays responses from the Fall and Spring, some questions were altered between semesters: Question 2 (Q2), Question 4 (Q4), and Question 9 (Q9). Question 7 (Q7) was only given in the Spring Semester. See **Appendix C** for a full list of questions and how they were altered.

Questions related to constraints (both in mammals and in birds) yielded the highest proportion of ultimate pattern descriptions (**Table X**). In response one, these two questions also yielded some of the highest proportions of teleological reasoning—the proportion of teleological reasoning related to constraints in mammals reduced in response two. In general, students reasoned that the persistence





(b) Proximate vs. Ultimate by Explanation vs. Description, R2

Figure 19: Mosaic plot displaying relationship between proximate versus ultimate descriptions and explanations. Descriptions and explanations are plotted along with their proportion of proximate and ultimate responses (levels) for response one (R1; a) and response two (R2; b). We observe that descriptions are most often at the ultimate level while explanations are most often at the proximate level. Shading of boxes reflects residuals: blue shading indicates more instances than expected if descriptions and explanations were random with respect to level; read indicates fewer instances than expected (p < 0.01).

of the four-limbed body plan or wings in penguins was solely the result of the adaptive significance of these traits and often invoked goal-directed processes. In both responses, the highest proportion of teleological reasoning resulted from the question relating to the loss of eyes in cave fish. A common justification was that fish "learned not to see" or "got rid of their eyes" because they were not necessary, which allowed them to "heighten their other senses." Ultimate pattern descriptions were invoked in fewer than half of responses in three questions: question five, relating to the maintenance of variation

(a) Proximate vs. Ultimate by Explanation vs. Description, R1



Figure 20: **Reasoning type by question.** Student responses may invoke strictly scientific reasoning, teleological reasoning, or some other kind of non-scientific (non-normative) reasoning. This plot displays reasoning type invoked by question in response one (R1; a) and response two (R2; b). Blue shading indicates more of that reasoning type than expected if reasoning were random with respect to question; red shading indicates less than expected (p < 0.01).

in populations; question eight, relating to the rise of beneficial traits in populations; and question six, relating to the diversity of flowering plants (only in response two). In these instances, proximate process explanations were more common and yielded significantly less teleological reasoning.

QUESTION	Response One	Response Two
Q1: Why all land mammals possess four limbs		
% Ultimate Pattern	90	93
% Teleological	58	38
Q3: Why penguins have wings if unable to fly		
% Ultimate Pattern	75	73
% Teleological	54	59
Q2: Why woody tissue evolved repeatedly across groups		
% Ultimate Pattern	67	76
% Teleological	31	41
Q10: Why there is no single perfectly adapted species, but millions of different ones		
% Ultimate Pattern	60	68
% Teleological	48	48
Q1: Why cave fish eventually lose their eyesight		
% Ultimate Pattern	57	57
% Teleological	73	62
Q7: Why some species are endemic and some cosmopolitan		
% Ultimate Pattern	55	59
% Teleological	13	15
Q4: Why some male birds of the same species maintain coat color polymorphism		
% Ultimate Pattern	51	16
% Teleological	13	8
Q6: Why flowering plants so diverse		
% Ultimate Pattern	50	35
% Teleological	28	26
Q5: Why natural selection does not bring all species to optimum		
% Ultimate Pattern	41	30
% Teleological	25	35
Q5: How beneficial traits arise in populations		
% Ultimate Pattern	40	24
% Teleological	47	22

TABLE X: **Percent ultimate pattern description and teleological response by question.** The most common responses were ultimate pattern descriptions overall, but this varied by individual question. Individual questions also invoked varying levels of teleological responses. Here we show the percent of ultimate pattern descriptions provided for each question, and the overall percent of teleological responses in response one and response two.



Figure 21: **Reasoning type by descriptions versus explanations**. Student responses may invoke strictly scientific reasoning, teleological reasoning, or some other kind of non-scientific (non-normative) reasoning. This plot displays reasoning type by descriptions versus explanations in response one (R1; a) and response two (R2; b). Blue shading indicates more of that reasoning type than expected if reasoning were random with respect to description versus explanation; red shading indicates less than expected (p < 0.01). We find that descriptions are disproportionately teleological, while explanations are disproportionately scientific.

# 5.7 Discussion

We find little work that examines students' propensity for providing descriptions versus explanations. This is surprising given that a major goal of science education is bridging the divide between purely descriptive knowledge and explanatory knowledge. Such a skill is especially vital in biological sciences and evolution in particular, where identifying patterns is a necessary step in providing robust causal explanations. We are not aware of any other framework that integrates these features, nor demonstrates the relationships between individual-level and species-level patterns and processes. Thus our framework provides unique insights into student thinking as it relates to adaptation, and a novel way to systematically assess student thinking in a broad and generalizable manner. We present preliminary data from two semesters that suggests that students overwhelmingly prefer ultimate pattern descriptions in response to questions related to adaptation. We also find, as many studies have documented before, a tendency toward the teleological. However, both of these observations appear to be predominantly dependent on the question being asked. Below we discuss some of the implications of our findings.

Teleological reasoning poses a consistent problem in the teaching of evolution, documented extensively in the literature (Barnes et al., 2017; Hammann & Nehm, 2020; Kampourakis, 2020; Kelemen, 2012; Lennox & Kampourakis, 2013; Tamir & Zohar, 1991). Cognitive origins of teleological reasoning are thought to arise for both domain-general and domain-specific reasons (for a full review, see Trommler & Hammann, 2020). Domain-general cognitive origins are thought to arise from human intuitions: dual-process models indicate that intuitive reasoning—as opposed to reflective reasoning occurs quickly and without voluntary control, leading to a host of documented biases. Domain-specific cognitive origins are less well-studied, but do suggest that adaptive traits in particular elicit this kind of reasoning. This is thought to arise for several reasons. First, human beings are inundated each day with artifacts that were designed for a purpose. Human-designed artifacts are likely the first and most consistent objects with which babies and children interact. This leads to a bias that *all* artifacts have a particular function or are designed for a particular reason. Second, adaptation is typically taught by noting the functional or adaptive significance of traits within a particular context. This emphasis exacerbates the teleological reasoning problem, because these types of explanations fit more neatly into students' intuitive reasoning related to the design of artifacts. The process of natural selection invites more complexity, which may make these types of explanations more challenging to fully internalize. Lastly, while largely unaddressed in the education literature, it seems that students are rarely taught how to provide appropriate explanations for the questions they are asked in an educational setting. We suspect this makes it especially challenging to provide them appropriately and in a consistent manner.

We find that formal instruction affects the overall pattern of responses very little. Students respond to questions in similar ways regardless of their major or how many earned credits of post-secondary biology. This result is similar to other studies that indicate formal instruction does little to address the most common misconceptions in student thinking regarding evolutionary theory (e.g., Coley & Tanner, 2015). Moreover, our result matches that of Riemeier (2009) almost exactly: overall, 55-57 percent of students invoke ultimate pattern descriptions. Of these, 46-49 percent invoke teleological reasoning. (Teleological reasoning makes up 40-36 percent of answers overall, reduced slightly in response two.) We find that while major, course, previously earned credits, or semester do not correlate strongly with response or reasoning, the questions themselves do. This fits with previous work demonstrating that novices are especially sensitive to item surface features. Instead of developing robust frameworks that provide explanatory power in many contexts, it seems that students form many superficial frameworks that provide explanatory power in a single context. For instance, students are routinely taught that sexual selection is common among birds using the classical example of the peacock. Based on our preliminary data, it appears that students—whenever the word "bird" appears in a question related to evolution—immediately invoke sexual selection, regardless of whether or not the question relates to sexual selection.

Indeed, questions that contain similar features of frequently-used classroom examples (e.g., the peacock and sexual selection) tend to elicit these examples as responses, regardless of whether or not the question is actually asking about the processes involved in those familiar examples. Such classroom examples are often invoked (e.g., the giraffe, the moth, the finches) as ultimate descriptions to explain why *other* organisms are also uniquely suited to their environments. Interestingly, when classroom examples relate to processes rather than patterns, these too get invoked: the process of sexual selection in birds or the process of pollination to explain the diversity of flowering plants. While non-normative ideas were common in the latter (many suggested that inter-specific mating during pollination was the reason for floral diversity), responses were notably more focused on process than in other questions. This suggests that students are tapping into common classroom examples and narratives, but perhaps have few examples of ultimate processes from which to draw upon. We suspect that the use of this framework as a guide to develop instructional methodologies may help with this possible imbalance, because it requires addressing patterns and processes at two levels. We expand on this more in **Chapter 6**.

We also observe that teleological reasoning is less common and scientific reasoning is more common when students invoke processes rather than describe patterns. The processes that students do invoke are largely proximate. Based on results from our semi-structured interviews, students appear more well-versed in proximate processes through their molecular and cell biology coursework. These processes also likely involve fewer complex reasoning skills, because they tend to occur in a linear and sequential manner—unlike natural selection which operates at multiple levels and includes emergent features. When asked for explanations of adaptation, all respondents first provided descriptions of organisms that are uniquely suited to their environment. When pressed for the process responsible for those observations, students struggled to provide coherent answers. All recognized that organisms cannot change their own DNA, nor can they decide the manner in which they want to evolve. But linking the proximate processes from molecular biology to changes in phenotypes over time presented an enormous barrier to understanding. We cannot help but muse that perhaps this observation relates to the fractured nature of the discipline itself—students take siloed courses in separate subdisciplines, and rarely is content integrated across such subdisciplines in a unified way.

Adaptation in textbooks commonly appears as examples of organisms' traits that are suited to their environments, and classroom teaching tends to also follow this pattern. Limited work investigates the extent to which educators discuss patterns versus processes within the context of class, but a few studies indicate that class content often focuses on low-level cognitive skills suggestive of descriptive knowledge or pattern recognition (e.g., Momsen et al., 2010). Thus, perhaps it is not surprising that students reason about adaptation in static ways given that (1) this method of reasoning is intuitive and established early in childhood (Inagaki & Hatano, 2002), and (2) textbooks and classroom examples tend to reinforce this thinking. Thus, not only is adaptation a particularly challenging concept about which to formulate true causal explanations (as many evolutionary biologists will agree), but science classes rarely attend to the features of true explanations within the context of science. Students are then on their own to integrate knowledge in a manner that allows for robust and cohesive explanatory power. The proposed framework is novel because it addresses the complexity of adaptation broadly and provides a clear means for demonstrating how to construct appropriate explanations. In biology, a discipline that on the whole is rarely linear and sequential, pattern recognition is an essential first step to providing causal justification. Students perform quite well at pattern recognition, providing a method of interacting with previous knowledge—we first begin with intuitive pattern recognition, but then work to show students how to link that pattern to an appropriate causal explanation. It is our job as educators to aid our students in building necessary skills within our disciplines. Thus, guiding students from the intuitive domain of descriptive knowledge to one of causal explanatory knowledge requires our support within the classroom.

### 5.7.1 Limitations and future work

This study shows preliminary data from one population of students during the Covid-19 pandemic. Instruction was entirely online during this academic year and may influence the results of the study. More work is needed to assess the validity of the questionnaire and the associated framework. More demographic data should be collected in tandem—including the use of validated instruments that measure understanding of natural selection—to determine how student construction of responses is related to their understanding generally. Lastly, the study should be repeated with an instructional intervention where the framework and its components are taught explicitly, providing the opportunity for students to become familiarized with the features of robust explanations and how to construct one in biological sciences. An intervention of this kind would provide valuable information on the utility of the framework and associated questionnaire. In general, frameworks provide scope through which we can view particular concepts and problem; they help organize and focus our knowledge. It is likely that many explanatory frameworks are required for a robust understanding of evolution by natural selection, some very broad and some quite specific. This broad framework provides a first step in helping students establish their own explanatory framework that integrates biological knowledge from two levels and distinguishes between patters and their generating processes. We suspect that aiding in the development of classroom materials may be its most powerful feature, discussed more in **Chapter 6**.

#### 5.8 Cited Literature

- Abrams, E., & Southerland, S. (2001). The how's and why's of biological change: How learners neglect physical mechanisms in their search for meaning. *International Journal of Science Education, 23*(12), 1271–1281.
- AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, *39*(10), 952–978.
- Barnes, M. E., Evans, E. M., Hazel, A., Brownell, S. E., & Nesse, R. M. (2017). Teleological reasoning, not acceptance of evolution, impacts students' ability to learn natural selection. *Evolution: Education and Outreach*, *10*(1), 1–12.
- Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of traitdependent speciation and extinction. *Systematic Biology*, *65*(4), 583–601.

Bechtel, W. (2011). Mechanism and biological explanation. Philosophy of Science, 78(4), 533-557.

- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421–441.
- Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, *27*(5), 415–427.
- Bradshaw, H., & Schemske, D. W. (2003). Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature*, *426*(6963), 176–178.
- Braillard, P.-A., & Malaterre, C. (2015). Explanation in biology: An introduction. In *Explanation in Biology* (pp. 1–28). Springer.
- Chi, M. T., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, *36*(1), 1–61.
- Coley, J. D., & Tanner, K. (2015). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. *CBELife Sciences Education*, *14*(1), ar8.
- Crespi, B. J. (2000). The evolution of maladaptation. *Heredity*, 84(6), 623–629.
- Cummins, C., & Remsen Jr, J. (1992). The importance of distinguishing ultimate from proximate causation in the teaching and learning of biology. *Proceedings of the Second International Conference for History and Philosophy of Science in Science Teaching*, *1*, 201–210.
- Demastes, S. S., Settlage Jr, J., & Good, R. (1995). Students' conceptions of natural selection and its role in evolution: Cases of replication and comparison. *Journal of Research in Science Teaching*, *32*(5), 535–550.

- Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher*, *35*(3), 125–129.
- Duncan, R. G. (2007). The role of domain-specific knowledge in generative reasoning about complicated multileveled phenomena. *Cognition and Instruction*, *25*(4), 271–336.
- Durbin, M. L., Lundy, K. E., Morrell, P. L., Torres-Martinez, C. L., & Clegg, M. T. (2003). Genes that determine flower color: The role of regulatory changes in the evolution of phenotypic adaptations. *Molecular Phylogenetics and Evolution*, *29*(3), 507–518.
- Evans, E. M. (2008). Conceptual change and evolutionary biology: A developmental analysis. *International Handbook of Research on Conceptual Change*, 263–294.
- Ferrari, M., & Chi, M. T. (1998). The nature of naïve explanations of natural selection. *International Journal of Science Education*, *20*(10), 1231–1256.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378.
- Forkmann, G., & Ruhnau, B. (1987). Distinct substrate specificity of dihydroflavonol 4-reductase from flowers of Petunia hybrida. *Zeitschrift Für Naturforschung C*, *42*(9-10), 1146–1148.
- Friendly, M. (2021). vcdExtra. https://CRAN.R-project.org/package=vcdExtra
- Gamer, M., Lemon, J., Gamer, M. M., Robinson, A., & Kendall's, W. (2012). Package "irr". Various Coefficients of Interrater Reliability and Agreement, 22.
- Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., & Igić, B. (2010). Species selection maintains self-incompatibility. *Science*, *330*(6003), 493–495.

- Greene Jr, E. D. (1990). The logic of university students' misunderstanding of natural selection. *Journal* of Research in Science Teaching, 27(9), 875–885.
- Gregory, T. R. (2009). Understanding natural selection: Essential concepts and common misconceptions. *Evolution: Education and Outreach*, *2*(2), 156–175.
- Gropengiesser, H., Harms, H., & Kattmann, U. (2016). *Fachdidaktik Biologie [Biology Education]* (Twelfth). Friedrich Verlag.
- Ha, M., Wei, X., Wang, J., Hou, D., & Nehm, R. H. (2019). Chinese pre-service biology teachers' evolutionary knowledge, reasoning patterns, and acceptance levels. *International Journal of Science Education*, *41*(5), 628–651.
- Habu, Y., Hisatomi, Y., & Iida, S. (1998). Molecular characterization of the mutable flaked allele for flower variegation in the common morning glory. *The Plant Journal*, *16*(3), 371–376.
- Hammann, M., & Nehm, R. H. (2020). *Teleology and evolution education: Introduction to the special issue*. Springer.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15*(2), 135–175.
- Inagaki, K., & Hatano, G. (2002). Young Children's Naive Thinking about the Biological World. Psychology Press.
- Kalinowski, S. T., Leonard, M. J., & Taper, M. L. (2016). Development and validation of the conceptual assessment of natural selection (CANS). *CBE–Life Sciences Education*, *15*(4), ar64.
- Kampourakis, K. (2013). Teaching about adaptation: Why evolutionary history matters. *Science & Education*, *22*(2), 173–188.

- Kampourakis, K. (2020). Students' "teleological misconceptions" in evolution education: Why the underlying design stance, not teleology per se, is the problem. *Evolution: Education and Outreach*, *13*(1), 1–12.
- Kampourakis, K., & Niebert, K. (2018). Explanation in Biology Education. In *Teaching Biology in Schools* (pp. 236–248). Routledge.
- Kampourakis, K., & Zogza, V. (2008). Students' intuitive explanations of the causes of homologies and adaptations. *Science & Education*, *17*(1), 27–47.
- Kelemen, D. (2012). Teleological minds. In *Evolution challenges: Integrating research and practice in teaching and learning about evolution*.
- Lennox, J. G., & Kampourakis, K. (2013). Biological teleology: The need for history. In *The Philosophy of Biology* (pp. 421–454). Springer.
- Maddison, W. P. (2006). Confounding asymmetries in evolutionary diversification and character change. *Evolution*, *60*(8), 1743–1746.
- Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, *56*(5), 701–710.
- Mayr, E. (1961). Cause and effect in biology. *Science*, *134*(3489), 1501–1506.
- Mayr, E. (1991). The ideological resistance to Darwin's theory of natural selection. *Proceedings of the American Philosophical Society*, 135(2), 123–139.
- Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2011). Recently formed polyploid plants diversify at lower rates. *Science*, *333*(6047), 1257–1257.
  Mazerolle, M. J., & Mazerolle, M. M. J. (2017). Package "AICcmodavg". *R Package*, *281*.

- Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE–Life Sciences Education*, 9(4), 435–440.
- Nehm, R. H., Beggrow, E. P., Opfer, J. E., & Ha, M. (2012). Reasoning about natural selection: Diagnosing contextual competency using the ACORNS instrument. *The American Biology Teacher*, *74*(2), 92–98.
- Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching*, 48(3), 237–256.
- Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, *57*(3), 263–272.
- Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, *45*(10), 1131–1160.
- Nesse, R. M. (2013). Tinbergen's four questions, organized: A response to Bateson and Laland. *Trends in Ecology & Evolution, 28*(12), 681–682.
- Pittendrigh, C. S. (1958). Adaptation, natural selection, and behavior. *Behavior and Evolution*, 390, 416.
- Potochnik, A. (2013). Biological explanation. In The Philosophy of Biology (pp. 49-65). Springer.
- Potochnik, A. (2020). What constitutes an explanation in biology? In *Philosophy of Science for Biologists* (pp. 21–35). Cambridge University Press.

- Ramsey, J., Bradshaw Jr, H., & Schemske, D. W. (2003). Components of reproductive isolation between the monkeyflowers Mimulus lewisii and M. Cardinalis (Phrymaceae). *Evolution; International Journal* of Organic Evolution, 57(7), 1520–1534.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
- Riemeier, T. (2009). Wie erklären Schüler biologische Phänomene?–Eine Analyse von Schülerbeschreibungen und–erklärungen. [How do students explain biological phenomena– An analysis of students' descriptions and explanations.]. In *Heterogenität erfassen–individuell fördern im Biologieunterricht* (pp. 71–83). Studienverlag.
- Robertshaw, D. (1971). The evolution of thermoregulatory sweating in man and animals. *International Journal of Biometeorology*, *15*(2), 263–267.
- Russ, R. S., Coffey, J. E., Hammer, D., & Hutchison, P. (2009). Making classroom assessment more accountable to scientific reasoning: A case for attending to mechanistic thinking. *Science Education*, *93*(5), 875–891.
- Salmon, W. C. (2006). Four decades of scientific explanation. University of Pittsburgh Press.
- Sánchez-Cabrera, M., Jiménez-López, F. J., Narbona, E., Arista, M., Ortiz, P. L., Romero-Campero, F. J., Ramanauskas, K., Igić, B., Fuller, A. A., & Whittall, J. B. (2021). Changes at a critical branchpoint in the anthocyanin biosynthetic pathway underlie the blue to orange flower color transition in *Lysimachia arvensis*. *Frontiers in Plant Science*, *12*, 247.

- Sbeglia, G. C., & Nehm, R. H. (2020). Illuminating the complexities of conflict with evolution: Validation of the scales of evolutionary conflict measure (SECM). *Evolution: Education and Outreach*, *13*(1), 1–22.
- Schemske, D. W. (2000). Understanding the origin of species. *Evolution*, 54(3), 1069–1073.
- Schofield, J. (2021). Towards a teleonomic philosophy of biology. In *A phenomenological revision of EE harris's dialectical holism* (pp. 207–258). Springer.
- Scott, E. E., Anderson, C. W., Mashood, K., Matz, R. L., Underwood, S. M., & Sawtelle, V. (2018). Developing an analytical framework to characterize student reasoning about complex processes. *CBE–Life Sciences Education*, *17*(3), ar49.
- Southerland, S. A., Abrams, E., Cummins, C. L., & Anzelmo, J. (2001). Understanding students' explanations of biological phenomena: Conceptual frameworks or p-prims? *Science Education*, *85*(4), 328–348.
- Speth, E. B., Shaw, N., Momsen, J., Reinagel, A., Le, P., Taqieddin, R., & Long, T. (2014). Introductory biology students' conceptual models and explanations of the origin of variation. *CBE–Life Sciences Education*, *13*(3), 529–539.
- Stearns, S. C. (2013). Natural selection, adaptation, and fitness: Overview. In *The Princeton Guide to Evolution* (pp. 193–199).
- Streisfeld, M. A., & Rausher, M. D. (2009). Genetic changes contributing to the parallel evolution of red floral pigmentation among Ipomoea species. *New Phytologist*, *183*(3), 751–763.
- Tamir, P., & Zohar, A. (1991). Anthropomorphism and teleology in reasoning about biological phenomena. *Science Education*, *75*(1), 57–67.

- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Für Tierpsychologie*, *20*(4), 410–433.
- Trommler, F., Gresch, H., & Hammann, M. (2018). Students' reasons for preferring teleological explanations. *International Journal of Science Education*, 40(2), 159–187.
- Trommler, F., & Hammann, M. (2020). The relationship between biological function and teleology: Implications for biology education. *Evolution: Education and Outreach*, *13*, 1–16.
- Vincent, T. L., & Brown, J. S. (2005). *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics*. Cambridge University Press.
- Ware, E. A., & Gelman, S. A. (2014). You get what you need: An examination of purpose-based inheritance reasoning in undergraduates, preschoolers, and biological experts. *Cognitive Science*, *38*(2), 197–243.

# **CHAPTER 6**

# UNIVERSITY STUDENTS' DESCRIPTIONS & EXPLANATIONS OF ADAPTATION. II. A FRAMEWORK FOR PEDAGOGICAL EXPLANATION.

# 6.1 Abstract

In **Chapter 5**, I present the development of a framework for systematically assessing student thinking as it relates to the concept of adaptation and the process of natural selection. I also present data analyzed with this framework from two semesters that includes nearly 600 student participants. Here, I present the utility of the proposed framework as a guide for informing instructional methodologies and developing educational materials. In the few empirical studies that examine students' propensity to provide descriptions versus explanations in relation to adaptive traits, students overwhelmingly prefer descriptions. Moreover, many of these descriptions are teleological in nature. This pattern appears remarkably resistant to formal instruction.

While an essential component of student understanding, pedagogical explanation is a poorly defined concept and an understudied area of education research. However, some work suggests that teachers may often provide descriptions in the place of explanations, or provide explanations with teleological or anthropocentric justifications. With especially challenging concepts like adaptation, providing simple and robust explanatory frameworks—without teleological or anthropocentric justification. I suggest that explicitly teaching frameworks like the one proposed will help students organize their biological knowledge appropriately, providing increased ability to develop expert-
like thinking that does not rely on non-normative ideas. Such frameworks provide tangible means for students to organize biological facts appropriately and in context, and aid in the development of metacognitive and higher-order reasoning skills.

### 6.2 Background

The science of life is notoriously complex. There are many levels of biological organization (molecular, cellular, tissue, whole organism, populations, species interactions, food webs, ecosystems) with complicated overlapping processes occurring within and between levels. In other disciplines, it may be possible to begin instruction with basic concepts and slowly build to more complexity in a linear fashion. For example, there is a natural sequence in mathematics that begins with counting objects and demonstrating what numbers represent. It is not necessary to disentangle algebraic expressions until a firm understanding of basic arithmetic is established. But where to begin with biology? How do we construct a sequence of instruction that must be linear in a discipline that is not? Even beginning, as we regularly do, with the very basic components necessary for life-macromolecules, water-we ignore the functional and evolutionary contexts within which such components exist. Students often feel that biological sciences is simply an exercise of memorizing structures and details. Without context, we fail to demonstrate to our students how to integrate biological knowledge appropriately and build robust frameworks that make sense of the discipline in a unified way. The story of life on our planet is perhaps the most beautiful and profound story we know. Why, then, when we teach biology, do we so often highlight fact after fact without linking those facts to the overarching story that gives them meaning? As Hutchinson noted with respect to life, "the ecological theater and evolutionary play" (Hutchinson,

1965); the ecological and evolutionary dynamics that govern life provides the explanatory thread that ultimately explains, if not describes, all levels of biological organization.

Building explanatory frameworks with facts in context is essential to the process of robust understanding, but studies repeatedly show that students face enormous challenges building such frameworks in the domain of biological sciences and in evolutionary theory in particular. The purpose of what follows is two-fold. First, I demonstrate the use of a framework as a pedagogical tool for explaining concepts in a manner that allows for appropriate integration of biological information. Second, I demonstrate that this framework also may aid in students' ability to distinguish between descriptions identifying patterns—and explanations, which provide detail regarding causal processes and may help reduce the propensity for teleological reasoning. These are the crucial skills students must develop to build their own powerful explanatory frameworks that are essential to scientific thinking.

## 6.2.1 Pedagogical explanation

Philosophers of science have been debating about the nature of scientific explanation for decades (Salmon, 2006). The pursuit of formal laws that dictate a unified model of scientific explanation has remained elusive, with biological sciences being an arena that is especially thorny (Braillard & Malaterre, 2015). Which is to say that the very thing that makes biology difficult to teach—complex and overlapping processes at many levels of biological organization—makes biology difficult to explain, even for biologists. This is an important and often unappreciated point in biology education (but see Kampourakis & Niebert, 2018): our students understand concepts through our explanations. But just how good *are* our biological explanations? What explanatory frameworks do we utilize to teach our students

complex concepts? How are we aiding in their ability to integrate what we teach them with other biological information?

Most philosophers agree that scientific explanations should provide relevant and compelling evidence, along with a justification that demonstrates causality (Potochnik, 2013). Of course, pedagogical explanations need not follow exactly this recipe. The purpose of a pedagogical explanation is not the same as the purpose of a scientific explanation: what makes a good pedagogical explanation is far more dependent on context. Scientific explanations aim to get at some aspect of Truth, while pedagogical explanations aim to bring students closer to the current normative understanding of a concept (Horwood, 1988). What is required for publication in a scientific journal is not the same as what is required to aid in students' scientific understanding. Thus what makes a good pedagogical explanation will depend heavily on the nuances of the topic and on the collection of students to which a concept is being explained. What is surprising is that despite the obvious centrality of explanation to the practice of pedagogy, relatively few empirical papers examine the nature of teachers' explanations (see Geelan, 2012 for review of this work) or attempt to define the features of good pedagogical explanations generally.

A notable exception is Treagust & Harrison (2000), who use the classic Richard Feynman lecture "Atoms in Motion" to identify characteristics of effective pedagogical explanation in the content domain of physics. Feynman is widely regarded as an expert science teacher and communicator (Sykes, 1995). There is one overarching concept and eight subconcepts identified in Feynman's lecture, each subconcept with a corresponding micro-explanation. A tapestry of analogies, metaphors (including anthropomorphic and teleological metaphors), and components of scientific evidence form a cohesive narrative that together explain the main concept: "that all things are made of atoms—little particles that move around in perpetual motion" (Treagust & Harrison, 2000, pg. 1166). The central point of the analysis is that good pedagogical explanations differ substantially from scientific ones. Mainly, effective teaching must interact with students' prior knowledge and must do so using more varied explanatory frameworks (e.g., metaphors).

Of course, a major challenge in engaging with students' prior knowledge is that this characteristic will differ across different groups of students. Thus effective pedagogical explanations are highly dependent on student backgrounds, pre-conceptions, biases, and perceived or actual knowledge base. As a result of this, ways of explaining are also quite varied, including extensive use of metaphors and analogies that must resonate and be accessible to the students. This presents complications: while anthropomorphic and teleological metaphors are powerful tools for interacting with intuitive knowledge, their continued use has the potential to reinforce reasoning patterns that will ultimately hinder students in forming scientific explanatory models. It is not clear at what point anthropomorphic and teleological explanations are traded for explanations of scientific processes—or even if educators ever explicitly address the changing nature of their explanations as students continue formal instruction. Moreover, explanations that rely on teleological or anthropocentric reasoning may do less damage in physics or chemistry, where it is easier to disentangle the action of physical and inanimate objects from a goal-oriented perspective. But living systems are different. Since natural selection promotes adaptations that enhance an organism's survivorship and/or reproduction, it makes sense that students see teleological and anthropocentric thinking as actual scientific explanations or sources of causal relationships. Even Aristotle was unable to conceptualize the existence of living creatures without purposeful design ("natural teleology", Aristotle, 1991).

#### 6.2.2 Descriptions and explanations in teaching

Various disciplines have their own perspectives on defining the terms "description" and "explanation." For example, in data analysis, descriptive analysis summarizes data in aggregate in an effort to determine overarching patterns, while explanatory analysis attempts to determine the causal factors responsible for the observed patterns (Blaikie, 2003). Most definitions tend to agree: (1) descriptions are pieces of information or statements of fact, and (2) explanations connect relevant pieces of information with a justification that demonstrates causality. Another way to express this is that descriptions provide information on *what* has taken place while explanations provide information on *why* it has taken place (similar to the typology outlined in Gropengiesser et al., 2016). In biology, although not necessarily formally defined as such, we often use the terms "pattern" and "process" when we discuss descriptions versus explanations.

Empirical research demonstrates how students often provide descriptions in the place of explanations (Riemeier, 2009, **Chapter 5** of this work). Yet several papers have pointed out that the distinction between these terms in classroom materials is often extremely fuzzy (Ebbers & Rowell, 2002; Horwood, 1988). The word "explain" is regularly used in place of the word "elaborate" (e.g., from Campbell Biological Sciences textbook 11e: "What do you think would happen if you decreased the pH of the intermembrane space of the mitochondrion? Explain your answer."), and the terms "explain" and "describe" may be used interchangeably. To complicate matters, the metaphors and analogies so often used in the classroom are generally themselves descriptions, not causal explanations. Through anecdotal evidence, discussion of what a good explanation is or how to construct one seems noticeably absent from nearly all undergraduate science curricula. Given these observations, it is wholly unsurprising that students are often unable to produce causal explanations. Often they are not provided with clear means for establishing this complex and important skill.

Indeed, developing causal explanations is difficult, requiring sustained practice and instruction. But this skill is absolutely essential if we aim to graduate students that possess robust scientific literacy skills. Without appropriate training in the nuances of scientific causal explanations, it may be difficult for our students to make sense of the explanations provided by science on the whole. Descriptions are pervasive because they are usually intuitive and easy to understand. Causal explanations are often deceptive and counter-intuitive—for many reasons, especially in the case of evolution where processes operating at the individual level produce emergent features at the population level (Chi et al., 2012). Thus a major goal of science education must include two features. First, we must impart students the ability to interpret scientific explanations for observed phenomena. At some point the utility of analogies and metaphors does not outweigh the need for students to develop coherent and expert-like explanatory frameworks.

### 6.2.3 Teaching adaptation

Adaptation is a concept that is particularly challenging to teach. To make matters worse, Kim & Nehm (2011) summarize decades of research that indicates American science teachers routinely hold antievolutionist beliefs that impact the teaching of evolution in schools. Out of nearly 1000 surveyed high school biology teachers, more than half reported that they do not devote more than ten hours of instruction *over an entire year* to evolutionary processes, with two percent that do not cover it at all (Berkman et al., 2008). In general, the fabric of evolutionary teaching in American K-12 schools

varies widely, with differences between states, school districts, and even between individual teachers in the same institution—highly dependent on school board or school administration policies, community members, and teachers' personal religious beliefs (Rutledge & Warden, 2000). There are still serious challenges to student conceptual understanding, even for teachers who espouse evolutionary theory and its centrality to the biological sciences, feel comfortable teaching it, and work in environments supportive of doing so.

Lawson et al. (2007) describe a general Piegetian sequence of reasoning ability with three stages of increasing complexity: (1) descriptive reasoning, (2) reasoning via causal agents that are perceptible (e.g., weight, heat), and (3) reasoning via causal agents that are not perceptible (e.g., theoretical entities like natural selection). Students that score lower on basic reasoning tests are more likely to harbor misconceptions related to natural selection and genetics—in fact, basic reasoning skills are more significantly and consistently correlated with misconceptions than mental capacity or verbal intelligence (Lawson & Thompson, 1988). These skills require substantial effort beyond that of memorization because understanding adaptation is highly contextual. Adaptations as *"heritable traits* that *maximize* an organism's expected *fitness* given the *ecological circumstances"* concatenate a number of fascinating yet challenging concepts and processes. Furthermore, even this direct definition may be viewed as arguable among experts in evolutionary biology. Thus developing robust reasoning skills is essential in helping students develop appropriate explanatory frameworks related to adaptation: it is not enough to simply define the concept and provide examples. Students must be shown how to *interpret* these examples within an appropriate framework. Empirical work that evaluates differences in problem solving between experts and novices demonstrates major differences in how these two groups understand and approach the problem solving activity (e.g., Chi et al., 1981). Only one study examines these differences in the domain of evolutionary biology. Nehm & Ridgway (2011) find that novices employ a rich diversity of strategies to approach problems and that approaches were often highly sensitive to item surface features (i.e., superficial details in problem like the particular mammal mentioned). In contrast, expert-like reasoning was far less sensitive to superficial details and far more consistently applied (i.e., greater explanatory coherence). In **Chapter 5**, I similarly find that students provide explanations that invoke mixtures of normative and non-normative ideas, metaphors, and analogies meaning that incorrect answers are rarely incorrect in exactly the same way. I find, similar to both Nehm & Ridgway (2011) and Kampourakis & Zogza (2008), that teleological explanations of adaptation are indeed the most common, but the proportion of teleological responses depends strongly on the question begin explained. This fits in line with previous work that demonstrates that student reasoning is highly dependent on the surface features of the question.

Another complicating factor are the inconsistencies in textbooks that define adaptation. In a cursory review of ten undergraduate textbooks, Kampourakis (2013) finds that definitions exist across three categories: historical, ahistorical, and process-based. Natural selection is routinely omitted when providing a definition of adaptation. Moreover, many biology courses do not discuss natural selection at all (e.g., cell biology). This aligns with our findings that students have trouble conceptualizing the process responsible for adaptation: that is, the process of natural selection is not clearly linked as the force of evolution that promotes adaptations. In semi-structured interviews (**Chapter 5**), students routinely explain adaptation with the descriptive phrase "the organism adapted to its environment" as if this is sufficient justification for the existence of adaptive traits. When pressed, many recognize that organisms do not spontaneously change their own DNA and that a process must occur that facilitates evolutionary change. But most students do not immediately recognize how to form these process-based explanations on their own.

### 6.2.4 Model of belief change

Models that characterize how learners change their beliefs throughout instruction include conceptual change models, dissonance theory, and constructivist models. Here we align, in general, with dualprocess models, a set of theoretical frameworks that describe how individuals reason and make decisions (Kahneman, 2011). Dual-process models implicate two types of information processing: (1) System 1 is fast, intuitive, high-capacity, and does not rely on working memory, and (2) System 2 is slow, reflective, low-capacity, and relies on working memory. These two systems have also been referred to using other terms—implicit and explicit; central and peripheral—but here we employ the terms first described by Stanovich & West (2000), which do not imply a specific relationship to consciousness. The idea behind these models is that most learners are operating with their System 1 intuitive reasoning, relying on heuristics to make decisions. Conversely, activating System 2 requires both motivation and ability (see Stanovich, 2009 for full review, including complications unique to dual-process theory). The implication for this work is that reasoning about adaptation in teleological ways is an intuitive and comfortable heuristic for most students (System 1), and overriding this belief requires sustained effort and explicit intervention in the classroom (System 2).

#### 6.3 Educational recommendations

I find previously that students overwhelmingly prefer descriptions to explanations. I also find a high proportion of teleological justifications across responses. These findings are unsurprising given the following general observations noted previously in the literature:

- Early reasoning skills are characterized by descriptions of nature (e.g., the first stage of Bloom's Taxonomy involves recognizing facts)
- 2. Students struggle with the more advanced reasoning skills that involve recognizing appropriate causal explanations (Bartov, 1978; Tamir, 1985; Yip, 2009)
- 3. Students struggle to construct their own scientific causal explanations, especially related to emergent or otherwise complex phenomena (Chi et al., 2012; Nehm et al., 2010)
- Students struggle to override natural intuitive reasoning patterns that are strongly teleological and need support to accomplish this task (Evans, 2008; Friedler et al., 1993; Inagaki & Hatano, 2002)
- Siloed undergraduate biology courses rarely discuss the elements of robust explanations or help students build robust and coherent explanatory frameworks for biological phenomena—especially if such activities require cross-disciplinary knowledge (Momsen et al., 2010; Nehm, 2019)

Students obviously need support in the classroom to traverse the great divide between novice-like and expert-like thinking. I do not know of any simple and visually represented frameworks developed for explaining adaptation. I suspect that a framework like this, akin to the Punnet Square in Mendelian genetics, will aid students in making sense of the complexity inherent in understanding adaptive traits. Not only does it provide a method of demonstrating how multiple layers of information are integrated, but it also provides an opportunity to address the difference between teleological and teleonomic reasoning in context. Moreover, strategies that focus on tasks like concept mapping and model building are known to aid students in organizing and integrating their knowledge on a particular topic (Novak, 2010).

Thus, this framework has utility as an active-learning activity in undergraduate biology courses following the introduction of the concept of adaptation. I developed a series of open-ended questions that relate to adaptation at the species-level in a variety of different organisms. A few of these questions may be used as examples to introduce the framework: (1) the difference between descriptions (patterns) and explanations (processes), and how processes provide causal justification for patterns, and (2) the two broad scales of biological hierarchy, the level of the individual and the level of the species. It may be emphasized that noting the adaptive value of a trait in its environment (students' commonly preferred justification) is simply a species-level pattern but does not otherwise demonstrate any kind of causality. Once the basics of the framework are established, students can work in pairs to fill in each quadrant of the grid for new questions. Specifics on the learning goals (LG) are discussed below.

#### 6.3.1 LG1: Distinguish between descriptions and explanations

Distinguishing between observed patterns and causal processes that may account for such patterns is a central exercise in science. In disciplines that examine the action of long-term historical processes evolutionary biology, paleogeology, cosmology—where direct observation of phenomena are not possible, scientists instead rely on the documentation of patterns to inform inductive or causal inference. Proving causality is not easy. Recent efforts take advantage of computational power to evaluate mathematical models that would otherwise be impossible to examine. To aid students in developing scientific thinking skills, it is essential that the ability to distinguish between patterns and generating processes is established.

A helpful way to elucidate these differences is to couch them in the more familiar terms of description and explanation, where a description answers that question "what?" and an explanation answers the question "why?" Students are trained from an early age through classroom materials and textbooks that the word "explain" tends to simply mean "describe more." What is often missing in science education is an explicit treatment of what the word "explain" means in a scientific context and the expected features of an explanation. Individuals in a population vary in heritable traits. That is a description, an observed pattern. But *why* is there variation in heritable traits? Because the process of DNA replication—while extremely accurate—produces occasional errors. In most single-celled organisms these errors are carried over during cell division; in sexually reproducing organisms they are carried over when produced in gametes. Thus the processes responsible for the production of new alleles can be linked explicitly to the pattern of observed intraspecific variation through the appropriate quadrants in the framework.

### 6.3.2 LG2: Identify processes that generate variation

Previous work demonstrates that students have trouble connecting the molecular processes that generate variation to the emergent features of adaptive traits at the population level (Speth et al., 2014). A major aspect of this involves linking changes in genes to changes in phenotypes. Indeed, in **Chapter 5** I find that 38% of students do not mention mutation when asked to explain how beneficial traits arise in a population of organisms. These findings together indicate that students struggle to link the processes covered in molecular and genetics courses to the phenotypic evolutionary change covered

in evolution and ecology courses. This link must be made explicit. How do errors in DNA replication (e.g., during meiosis or binary fission) account for the appearance of new phenotypes? What is the difference between novel gene sequences and novel genotypes? How does crossing-over and independent assortment relate to the production of novel genotypes? As a proximate pattern, organisms are constructed differently: they tend to have different DNA. But what processes are responsible for that observed variation in individual organisms? The first individual-level processes are molecular, including mutation (all organisms), recombination (sexually reproducing organisms), or horizontal gene transfer.

Once the origin of variation is properly categorized, the variation will play out in the survival and reproduction of individual organisms. Sexual reproduction is another individual-level process that can produce novel genotypes among sexually reproducing organisms. In textbooks, "sexual selection" is often represented as a process distinct from natural selection. In reality, sexual selection represents one facet of natural selection: differential reproduction. Of course it is an interesting feature that selection does not always maximize fitness in both survivorship and fecundity. Adaptations that may increase the ability to find mates may indeed hinder survival in some cases. But while these two facets may have different effects on an organisms' overall fitness, fitness is still most certainly composed of both. Separating these facets as if peacocks are not subject to "regular" natural selection likely makes it challenging to form a robust understanding of natural selection that accounts for the potentially complex interplay between survival and reproduction.

In the previous chapter, I find that nearly a third of students (31%) invoke sexual selection in a question that mentions birds—a question that asks about an intraspecific male feather color polymorphism that is not related to sexual selection. Students are likely applying a familiar classroom example—sexual selection in birds—to a question that mentions birds because of the high sensitivity of novices to item surface features. Which is to say that we are not helping students recognize how the process of natural selection acts on all organisms in the same way, despite the fact that the outcome (observed pattern) may be different in different contexts. There is not a different kind of natural selection that operates on birds. Highlighting examples that demonstrate various features of selection without making it clear what connects *every example* likely serves to exacerbate sensitivity to item surface features. By using the framework for varying adaptations in varying contexts, students may start to get a sense of how contextdependency interacts with the basic features of heritable differences in survival and reproduction upon which natural selection is based.

### 6.3.3 LG3: Explain population-level features

Adaptation is most often taught by example. But examples are not processes. A common misconception among the general public is that individual organisms evolve. For example, in the previous chapter, many explain trait loss with the justification that an individual organism will lose unnecessary or unused traits (e.g., "If you do not use something the body will not maintain it, like if you don't exercise you will lose your muscles"). The problem with examples of organisms uniquely suited to their environment is that it leads to a disconnect between individuals that are not evolving and populations that are, making it challenging to override this initial misconception. Such examples rarely emphasize processes and tend to focus on population-level patterns of adaptation. Moreover, separate emphasis on evolutionary processes at the micro-level may obscure processes occurring at or above the population level that lead to evolutionary change. It also confounds phenotypic and behavioral plasticity in individual organisms versus the change in heritable characteristics of a population over time. There are many factors that influence the evolutionary pathways available to organisms. Of course, natural selection may only act upon existing variation. But ecological factors like changing optima or environmental heterogeneity also influence the action of natural selection and the resultant adaptations. The features or strategies of other organisms may also influence this landscape. What this means is that the process of natural selection—while the same process—may produce varied observable patterns when selection is directional, stabilizing, or disruptive. Students tend to assume that the process of natural selection is one that moves species to an optimum, that all selection is directional. This belief makes understanding frequency-dependent selection and the patterns it produces especially challenging. By placing various examples within the context of the framework, we may demonstrate more clearly how the exact same process may produce apparently different results.

### 6.3.4 LG4: Distinguish between teleonomic and teleological reasoning

The difference between teleological and teleonomic justification has long been recognized among evolutionary biologists, but the education world is more recently addressing this important distinction (Kampourakis, 2020; Trommler & Hammann, 2020). (Many earlier works recognize the challenges teleological pose for biology education, but do not distinguish between teleological justification and teleonomic justification.) While recognizing the adaptive value of a trait is not an appropriate causal justification for the existence of the trait, it is an important step in forming a hypothesis about the action of natural selection over time. Which is to say that recognizing species-level patterns of adaptive features is a major aspect of the study of evolutionary biology. Today, evolutionary biologists collect data on observed traits and then use computational modeling tools to infer the evolutionary history of such traits. Thus recognizing teleonomic patterns at the species level is an essential component of forming a robust understanding of natural selection. However, the methods that we use to distinguish the recognition of a pattern versus providing the pattern as a causal justification are crucial.

In the previous chapter, students overwhelmingly preferred describing species-level patterns of adaptation in response to causal questions. But in many cases these observations were not incorrect in and of themselves, they were only incorrect as causal justifications. These responses tended to be teleonomic observations used in place of explanations: overall, about 40% of ultimate pattern responses. The other 60% of ultimate pattern responses instead used teleological justifications that invoked either (1) most commonly, the internal motivation of the organism (e.g., "penguins needed to swim so they made their wings like flippers"), or (2) a goal-oriented outcome (e.g., "wood evolved because humans need it for shelter"). Previous work indicates that students are quickly able to make the distinction between teleological and teleonomic justifications when such a distinction is demonstrated clearly in the classroom (Galli et al., 2020). For instance, students can quickly see how the proto-penguins' behavior of becoming more aquatic made flipper-like wings fitness enhancing. Hence "internal motivation" becomes a shorthand for the activities of an organism selecting for traits suited to that lifestyle. And when presented with height in trees as an adaptation for competing with other plants for light, students grasp that humans having wood is an unintended consequence. Thus it seems that developing the ability is not difficult with explicit intervention in the classroom.

Explanation of the ultimate species-level pattern within the context of the proposed framework allows educators to distinguish between appropriate (teleonomic) observations about traits, versus inappropriate goal-oriented (teleological) justifications regarding the existence of such traits. First, the ability to distinguish patterns and processes in general should aid in reducing teleological justifications as explanations. If students can recognize that identifying the adaptive value of a trait is simply a pattern, then it makes clearer why this response does not provide sufficient explanation. Second, students generally recognize that an organism cannot change its own DNA—but this fact from molecular biology somehow gets disconnected from the process of natural selection. Putting facts into context within the structure of the framework may also reduce motive- or goal-oriented reasoning because such explanations do not fit biologically into its structure.

## 6.4 Final thoughts

What is clear from these studies is that integrating lines of reasoning across biological hierarchy to form consistent and coherent explanations is a major challenge. Conceptual change theory suggests that students need the opportunity to reform intuitive thinking patterns over time with repeated intervention. This is only possible if educators explicitly address intuitive reasoning in the classroom and demonstrate strategies for overriding these thinking patterns consistently across courses. Forming robust explanatory frameworks requires frequent practice, and, as previous work indicates, there are many stages between novice and expert understanding. The framework developed in the previous chapter provides guidance for instructional methodologies and classroom activities that aid in students' understanding of natural selection and adaptation. Explicitly teaching the framework and using it as an active exercise aids in integrating knowledge at many levels, identifying appropriate explanations, and showing the relationship between patterns and generating processes. Ultimately this empowers students to begin forming their own coherent explanatory frameworks.

#### 6.5 Cited Literature

- Aristotle. (1991). *The Complete Works of Aristotle: The revised Oxford translation* (J. Barnes, Ed.; Vol.1). Princeton University Press.
- Bartov, H. (1978). Can students be taught to distinguish between teleological and causal explanations?. Journal of Research in Science Teaching, 15(6), 567–572.
- Berkman, M. B., Pacheco, J. S., & Plutzer, E. (2008). Evolution and creationism in America's classrooms: A national portrait. *PLoS Biology*, 6(5), e124.
- Blaikie, N. (2003). Analyzing quantitative data: From description to explanation. Sage.
- Braillard, P.-A., & Malaterre, C. (2015). Explanation in biology: An introduction. In Explanation in Biology (pp. 1–28). Springer.
- Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152.
- Chi, M. T., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, *36*(1), 1–61.
- Ebbers, M., & Rowell, P. (2002). Description is not enough: Scaffolding children's explanations. *Primary Science Review*, *74*, 10–13.
- Evans, E. M. (2008). Conceptual change and evolutionary biology: A developmental analysis. *International Handbook of Research on Conceptual Change*, 263–294.
- Friedler, Y., Zohar, A., & Tamir, P. (1993). The effect of age and of learning on the ability to distinguish between anthropomorphic and teleological explanations. *International Journal of Science Education*, *15*(4), 439–443.

- Galli, L. G., Pérez, G., & Galindo, A. A. G. (2020). The self-regulation of teleological thinking in natural selection learning. *Evolution: Education and Outreach*, *13*(1), 1–16.
- Geelan, D. (2012). Teacher explanations. In *Second International Handbook of Science Education* (pp. 987–999).
- Gropengiesser, H., Harms, H., & Kattmann, U. (2016). *Fachdidaktik Biologie [Biology Education]* (Twelfth). Friedrich Verlag.
- Horwood, R. H. (1988). Explanation and description in science teaching. *Science Education*, *72*(1), 41–49.
- Hutchinson, G. E. (1965). The ecological theater and the evolutionary play. Yale University Press.
- Inagaki, K., & Hatano, G. (2002). Young Children's Naive Thinking about the Biological World. Psychology Press.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kampourakis, K. (2013). Teaching about adaptation: Why evolutionary history matters. *Science & Education*, *22*(2), 173–188.
- Kampourakis, K. (2020). Students' "teleological misconceptions" in evolution education: Why the underlying design stance, not teleology per se, is the problem. *Evolution: Education and Outreach*, *13*(1), 1–12.
- Kampourakis, K., & Niebert, K. (2018). Explanation in Biology Education. In *Teaching Biology in Schools* (pp. 236–248). Routledge.
- Kampourakis, K., & Zogza, V. (2008). Students' intuitive explanations of the causes of homologies and adaptations. *Science & Education*, *17*(1), 27–47.

- Kim, S. Y., & Nehm, R. H. (2011). A cross-cultural comparison of Korean and American science teachers' views of evolution and the nature of science. *International Journal of Science Education*, 33(2), 197– 227.
- Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching:* 44(5), 706–724.
- Lawson, A. E., & Thompson, L. D. (1988). Formal reasoning ability and misconceptions concerning genetics and natural selection. *Journal of Research in Science Teaching*, *25*(9), 733–746.
- Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE–Life Sciences Education*, 9(4), 435–440.
- Nehm, R. H. (2019). Biology education research: Building integrative frameworks for teaching and learning about living systems. *Disciplinary and Interdisciplinary Science Education Research*, 1(1), 1– 18.
- Nehm, R. H., Rector, M. A., & Ha, M. (2010). "Force-talk" in evolutionary explanation: Metaphors and misconceptions. *Evolution: Education and Outreach*, *3*(4), 605.
- Nehm, R. H., & Ridgway, J. (2011). What do experts and novices "see" in evolutionary problems? *Evolution: Education and Outreach*, *4*(4), 666–679.
- Novak, J. D. (2010). Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations. Routledge.

Potochnik, A. (2013). Biological explanation. In The Philosophy of Biology (pp. 49-65). Springer.

- Riemeier, T. (2009). Wie erklären Schüler biologische Phänomene?–Eine Analyse von Schülerbeschreibungen und-erklärungen. [How do students explain biological phenomena– An analysis of students' descriptions and explanations.]. In *Heterogenität erfassen-individuell fördern im Biologieunterricht* (pp. 71–83). Studienverlag.
- Rutledge, M. L., & Warden, M. A. (2000). Evolutionary theory, the nature of science & high school biology teachers: Critical relationships. *The American Biology Teacher*, 23–31.

Salmon, W. C. (2006). Four decades of scientific explanation. University of Pittsburgh Press.

- Speth, E. B., Shaw, N., Momsen, J., Reinagel, A., Le, P., Taqieddin, R., & Long, T. (2014). Introductory biology students' conceptual models and explanations of the origin of variation. *CBE–Life Sciences Education*, *13*(3), 529–539.
- Stanovich, K. E. (2009). *Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory?*
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(5), 645–665.
- Sykes, C. (1995). Preface. In *No ordinary genius: The illustrated Richard Feynman* (pp. 9–12). WW Norton & Company.
- Tamir, P. (1985). Causality and teleology in high school biology. *Research in Science & Technological Education*, *3*(1), 19–28.
- Treagust, D. F., & Harrison, A. G. (2000). In search of explanatory frameworks: An analysis of Richard Feynman's lecture "Atoms in motion". *International Journal of Science Education*, *22*(11), 1157–1170.

- Trommler, F., & Hammann, M. (2020). The relationship between biological function and teleology: Implications for biology education. *Evolution: Education and Outreach*, *13*, 1–16.
- Yip, C.-W. (2009). Causal and teleological explanations in biology. *Journal of Biological Education*, *43*(4), 149–151.

APPENDICES

## Appendix A

## APPENDIX A

#### A.0.1 Excerpt regarding republishing from publisher website

#### **Republish or Adapt Excerpts or Entire Article**

You have the nonexclusive right to republish your article, in whole or in part, in a form identical to the VoR or any previous version or as a new adaptation or derivative, in any subsequent book, article, or other scholarly work of which you are an author or an editor, including your dissertation, provided that appropriate credit is given to the journal as follows: If the article has not yet been published, include a statement specifying the article's status, date, and journal name. "For example: Submitted (or Accepted) for publication to (by) Journal Name on MM/DD/YYYY." Include the article DOI if available.

From the University of Chicago Press website: https://www.journals.uchicago.edu/cont/
jrnl\_rights#republish

### A.0.2 Detailed synonymy results

**Breeding system dataset**. All species designations classified as "Accepted" were kept (n = 1067). Species designations classified as "Synonym", "Illegitimate", "Invalid", "No opinion", "Misapplied", or "Rejected name" were replaced with Accepted names when available (n = 200) or removed from the datasets (n = 57). We also removed species no longer recognized in the family Fabaceae (n = 5). Some entries in our trait dataset did not have species epithets (e.g., *Acacia sp.*). In these instances genera were synonymized only.

**Chromosome counts dataset**. All species designations classified as "Accepted" were kept (n = 4185). Species designations classified as "Synonym", "Illegitimate", "Invalid", "No opinion", "Misapplied", or "Rejected name" were replaced with Accepted names when available (n = 514) or removed from the datasets (n = 248). We also removed species no longer recognized in the family Fabaceae (n = 20). Species designations that yielded no results (n = 10) were also removed from the dataset.

**Phylogenetic dataset.** All species designations classified as "Accepted" were kept (n = 3530). Species designations classified as "Synonym", "Illegitimate", "Invalid", "No opinion", "Misapplied", or "Rejected name" were replaced with Accepted names when available (n = 576) or removed from the datasets (n = 187). We also removed species no longer recognized in the family Fabaceae (n = 5). Species designations that yielded no results (n = 70) were also removed from the dataset.

#### A.0.3 List of legume genera lacking breeding system information in our collection

Abarema, Acaciella, Achyronia, Acosmium, Acrocarpus, Adenocarpus, Adenolobus, Afgekia, Afzelia, Aganope, Airyantha, Akschindlium, Alantsilodendron, Aldina, Alexa, Alhagi, Alistilus, Amblygonocarpus, Amburana, Amerimnon, Amherstia, Amicia, Ammodendron, Amoria, Amphimas, Anarthrophyllum, Angylocalyx, Anila, Antheroporum, Anthonotha, Aotus, Apalatoa, Aphanocalyx, Apoplanesia, Apuleia, Aragallus, Archidendron, Aspalathus, Astracantha, Augouardia, Austrosteenisia, Baikiaea, Balsamocarpon, Baphia, Baphiopsis, Barnebydendron, Barnebyella, Batesia, Baudouinia, Berlinia, Bikinia, Bionia, Bituminaria, Blanchetiodendron, Bobgunnia, Bobrovia, Bocoa, Bolusanthus, Bolusia, Bossiaea, Bowringia, Brachystegia, Brandzeia, Brasilettia, Brenierea, Brodriguesia, Brongniartia, Brownea, Browneopsis, Brya, Bryaspis, Burkea, Burtonia, Bussea, Cadia, Calicotome, Callerya, Calliandropsis, Calophaca, Calpocalyx, Calpurnia, Camoensia, Camptosema, Campylotropis, Candolleodendron, Cascaronia, Castanospermum, Cathormion, Cedrelinga, Cenostigma, Centrolobium, Chesneya, Chidlowia, Chloroleucon, Christia, Cladrastis, Cleobulia, Cochlianthus, Codariocalyx, Cojoba, Cologania, Colvillea, Conzattia, Copaiba, Corallodendron, Cordeauxia, Cordyla, Corethrodendron, Coulteria, Coursetia, Craibia, Craspedolobium, Crudia, Cyamopsis, Cyathostegia, Cyclolobium, Cylicodiscus, Cylindrokelupha, Cymbosema, Cynometra, Daniellia, Decorsea, Dequelia, Delonix, Dendrolobium, Detarium, Dialium, Dichilus, Dichrostachys, Dicorynia, Dicymbe, Didelotia, Dinizia, Diphysa, Diplotropis, Diptychandra, Distemonanthus, Disynstemon, Dolichopsis, Dorycnium, Dumasia, Dunbaria, Duparquetia, Dupuva, Dussia, Ebenopsis, Ebenus, Ecuadendron, Elephantorrhiza, Eligmocarpus, Elizabetha, Emerus, Endertia, Englerodendron, Entada, Entadopsis, Eperua, Erophaca, Errazurizia, Ervum, Etaballia, Eurypetalum, Exostyles, Eysenhardtia, Feuilleea, Fiebrigiella, Fillaeopsis, Fissicalyx, Fordia, Gagnebina, Galega, Genistidium, Gilbertiodendron, Gilletiodendron, Glycyrrhiza, Goniorrhachis, Grazielodendron, Griffonia, Gueldenstaedtia, Guianodendron, Guibourtia, Guilandina, Haematoxylum, Halimodendron, Hammatolobium, Haplormosia, Hardwickia, Harleyodendron, Harpalyce, Havardia, Heinekenia, Hesperothamnus, Heteroflorum, Heterostemon, Hoffmannseggia, Hoita, Holocalyx, Humularia, Hybosema, Hydrochorea, Hylodendron, Hylodesmum, Hymenolobium, Hymenostegia, Hypocalyptus, Icuria, Indigastrum, Inocarpus, Isoberlinia, Isotropis, Jacqueshuberia, Julbernardia, Kanaloa, Kotschya, Kunstleria, Lackeya, Ladeania, Lamprolobium, Lebeckia, Lebruniodendron, Lennea, Leonardendron, Leonardoxa, Leptoderris, Leptolobium, Lessertia, Leucochloron, Leucomphalos, Librevillea, Liparia, Loesenera, Lophocarpinia, Luetzelburgia, Lysidice, Lysiphyllum, Maackia, Machaerium, Macrolobium, Macrosamanea, Maniltoa, Marina, Mariosousa, Martiodendron, Melanoxylon, Melolobium, Mendoravia, Microberlinia, Microcharis, Microlobius, Mildbraediodendron, Millettia, Mimozyganthus, Moldenhawera, Monopteryx, Mora, Moullava, Muellera, Mundulea, Myrocarpus, Myroxylon, Neochevalierodendron, Neorautanenia, Neorudolphia, Nesphostylis, Newtonia, Nissolia, Normandiodendron, Oddoniodendron, Ophrestia, Orbexilum, Oreophysa, Ormocarpopsis, Ormocarpum, Ormosia, Orphanodendron, Ostryocarpus, Otoptera, Pachyelasma, Painteria, Paloue, Panurea, Paramachaerium, Paramacrolobium, Parapiptadenia, Pararchidendron, Paraserianthes, Parochetus, Parryella, Pearsonia, Pediomelum, Peltogyne, Peltophorum, Periandra, Pericopsis, Petaladenium, Peteria, Phanera, Philenoptera, Phyllodium, Phylloxylon, Physostigma, Pictetia, Piliostigma, Piptadeniastrum, Piptadeniopsis, Piptanthus, Pityrocarpa, Placolobium, Plagiocarpus, Plagiosiphon, Plathymenia, Platycyamus, Platymiscium, Podalyria, Poecilanthe, Poiretia, Poitea, Polystemonanthus, Pomaria, Pongamia, Pongamiopsis, Prioria, Prosopidastrum, Pseudoprosopis, Pseudosamanea, Pseudovigna, Psoralea, Psorothamnus, Pterogyne, Pterolobium, Ptycholobium, Rafnia, Ramirezella, Recordoxylon, Rehsonia, Rhodopis, Rhynchotropis, Riedeliella, Robinia, Robynsiophyton, Rupertia, Salwee-

nia, Saraca, Sartoria, Schleinitzia, Schnella, Schotia, Scorodophloeus, Securigera, Sellocharis, Shuteria, Sindora, Sindoropsis, Smirnowia, Smithia, Spartium, Spathionema, Spatholobus, Sphaerophysa, Sphenostylis, Sphinctospermum, Sphinga, Spiesia, Spirotropis, Stachyothyrsus, Stemonocoleus, Stenodrepanum, Stirtonanthus, Stizolobium, Storckiella, Strophostyles, Stryphnodendron, Stuhlmannia, Styphnolobium, Sweetia, Tadehagi, Talbotiella, Taralea, Taverniera, Templetonia, Teramnus, Tessmannia, Tetraberlinia, Tetrapleura, Thailentadopsis, Tibetia, Tipuana, Trischidium, Umtiza, Uribea, Vandasina, Vatairea, Vataireopsis, Vatovaea, Vavilovia, Viguieranthus, Virgilia, Vouacapoua, Wajira, Weberbauerella, Wiborgia, Wisteria, Xanthocercis, Xerocladia, Xeroderris, Xylia, Zapoteca, Zenia, Zenkerella, Zollernia

### A.0.4 Collected works for Fabaceae breeding system database

- Abberton, M. T. (2007). Interspecific hybridization in the genus *Trifolium*. *Plant Breeding*, 126(4), 337–342.
- Agbagwa, I. O., & Obute, G. C. (2007). Breeding system and pollination ecology of two *Abrus* species (Fabaceae) from tropical West Africa. *Acta Botanica Croatica*, 66(2), 205–216.
- Agbagwa, I. O., & Okoli, B. E. (2005). Chromosome number and cytomorphological characterization of a polyploid *Abrus*. *African Journal of Biotechnology*, *4*(7), 607–610.
- Agostini, K. (2008). Ecologia da reprodução de duas especies de Mucuna (Leguminosae, Faboideae, Phaseoleae): Embriologia, citogenetica e variabilidade genetica-do litoral norte de São Paulo [PhD Thesis]. Universidade Estadual de Campinas.
- Aguiar, L. F. R. (1992). Fenologia, sistema de reproducão, ecologia da polinização e dispersão de Senna alata (Caesalpinioideae, Leguminosae) [PhD Thesis]. Universidade Estadual de Campinas.
- Aizen, M. A., & Feinsinger, P. (1994). Forest fragmentation, pollination, and plant reproduction in a Chaco dry forest, Argentina. *Ecology*, 75(2), 330–351.
- Allphin, L., Brian, N., & Matheson, T. (2005). Reproductive success and genetic divergence among varieties of the rare and endangered *Astragalus cremnophylax* (Fabaceae) from Arizona, USA. *Conservation Genetics*, 6(5), 803–821.
- Almeida, N. M., Bezerra, T. T., Oliveira, C. R. S., Novo, R. R., Siqueira-Filho, J. A., Oliveira, P. E., & Castro, C. C. (2015). Breeding systems of enantiostylous Cassiinae species (Fabaceae, Caesalpinioideae). Flora-Morphology, Distribution, Functional Ecology of Plants, 215, 9–15.
- Almeida, N. M., Castro, C. C., Leite, A. V., Novo, R. R., & Machado, I. C. (2013). Floral polymorphism in *Chamaecrista flexuosa* (Fabaceae-Caesalpinioideae): A possible case of atypical enantiostyly? *Annals* of Botany, 112(6), 1117–1123.
- Ariwaodo, J. O., Harry-Asobara, J. L., & others. (2015). Preliminary investigation on flowering and fruiting pattern in a plantation grown *Afzelia africana* Sm. Stand in Umuahia, Nigeria. *American Journal of Plant Sciences*, 1(2), 219.
- Arroyo, M. K. (1981). Breeding systems and pollination biology in Leguminosae. In R. Polhill & P. Raven (Eds.), Advances in Legume Systematics, Part 2 (pp. 723–769). Royal Botanic Gardens, Kew.
- Arroyo, M. T. K., & Uslar, P. (1993). Breeding systems in a temperate mediterranean-type climate montane sclerophyllous forest in central Chile. *Botanical Journal of the Linnean Society*, 111(1), 83–102.
- Artyukova, E., Kholina, A., Kozyrenko, M., & Zhuravlev, Y. N. (2004). Analysis of genetic variation in rare endemic species Oxytropis chankaensis Jurtz. (Fabaceae) using RAPD markers. Russian Journal of Genetics, 40(7), 710–716.

- Atlan, A., Schermann-Legionnet, A., Udo, N., & Tarayre, M. (2015). Self-incompatibility in *Ulex europaeus*: Variations in native and invaded regions. *International Journal of Plant Sciences*, 176(6), 515–524.
- Australian Native Plants Society. (2010). In Acacia ulicifolia. http://anpsa.org.au/a-uli.html
- Baker, H., & Harris, B. (1957). The pollination of *Parkia* by bats and its attendant evolutionary problems. *Evolution*, 449–460.
- Baker, P. J. (2009). *Koa* (Acacia koa) *ecology and silviculture* (Vol. 211). United States Department of Agriculture, Forest Service,
- Barbosa, R. I., Ferreira, E. J. G., & Castellón, E. G. (1997). Homem, ambiente e ecologia no estado de roraima. Inpa.
- Barker, H. G. (1970). Two cases of bat pollination in Central America. *Revista de Biologia Tropical*, *17*(2), 187–197.
- Barros, E. C. de O., Webber, A. C., & Machado, I. C. (2013). Limitação de polinizadores e mecanismo de autoincompatibilidade de ação tardia como causas da baixa formação de frutos em duas espécies simpátricas de *Inga* (Fabaceae-Mimosoideae) na Amazônia Central. *Rodriguésia*, 64(1), 37–47.
- Baskin, C. C., Baskin, J. M., & Quarterman, E. (1972). Observations on the ecology of Astragalus tennesseensis. American Midland Naturalist, 88(1), 167–182.
- Bawa, K. S. (1974). Breeding systems of tree species of a lowland tropical community. *Evolution*, 28(1), 85–92.
- Bawa, K. S., Perry, D. R., & Beach, J. H. (1985). Reproductive biology of tropical lowland rain forest trees.
   I. Sexual systems and incompatibility mechanismsexual systems and incompatibility mechanisms. *American Journal of Botany*, 72(3), 331–345.
- Belov, M. (2009). Chileflora. In *Description and Images of* Adesmia mucronata. http://www.chileflora .com/Florachilena/FloraEnglish/HighResPages/EH0701.htm
- Bena, G., Lejeune, B., Prosperi, J.-M., & Olivieri, I. (1998). Molecular phylogenetic approach for studying life-history evolution: The ambiguous example of the genus *Medicago* L. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1401), 1141–1151.
- Bencie, R. (1997). *Genetic variation and inbreeding depression in the rare California endemic,* Astragalus agnicidus (*Leguminosae*) [PhD thesis]. Humboldt State University.
- Ben-Zeev, N., & Zohary, D. (1973). Species relations in the genus *Pisum* L. *Isreal Journal of Botany*, *22*, 73–91.
- Bergamo, P. J., & Sazima, M. (2018). Differential outcomes of nectar robbing on the reproductive success of a melittophilous and an ornithophilous species. *International Journal of Plant Sciences*, 179(3), 192–197.
- Bernhardt, P., Kenrick, J., & Knox, R. (1984). Pollination biology and the breeding system of Acacia retinodes (Leguminosae, Mimosoideae). Annals of the Missouri Botanical Garden, 71(1), 17–29.
- Bessega, C., Ferreyra, L., Juliof, N., Montoya, S., Saidmanj, B., & Vilardij, J. C. (2000). Mating system parameters in species of genus *Prosopis* (Leguminosae). *Hereditas*, *132*(1), 19–27.
- Biswas, M., & Dana, S. (1976). Phaseolus aconitifolius × p. Trilobatus. Indian Journal of Genetics and Plant Breeding, 36, 125–131.
- Boaler, S. (1966). Ecology of Pterocarpus angolensis DC in Tanzania.

- Boe, A., & Fluharty, K. (1993). Reproductive biology of a Canada milk-vetch population from eastern South Dakota. *Prairie Naturalist*, *25*, 65–65.
- Bogdan, A., & others. (1977). *Tropical pasture and fodder plants*. Longman.
- Bohart, G. E. (1960). Insect pollination of forage legumes. *Bee World*, 41(4), 85–97.
- Borges, L. A., Machado, I. C., & Lopes, A. V. (2017). Bee pollination and evidence of substitutive nectary in *Anadenanthera colubrina* L. (Leguminosae-Mimosoideae). *Arthropod-Plant Interactions*, *11*, 1–9.
- Borges, L. A., Sobrinho, M. S., & Lopes, A. V. (2009). Phenology, pollination, and breeding system of the threatened tree *Caesalpinia echinata* Lam. (Fabaceae), and a review of studies on the reproductive biology in the genus. *Flora*, 204(2), 111–130.
- Borges, L. A., Souza, L. G. R., Guerra, M., Machado, I. C., Lewis, G. P., & Lopes, A. V. (2012). Reproductive isolation between diploid and tetraploid cytotypes of *Libidibia ferrea* (= *Caesalpinia ferrea*) (Leguminosae): Ecological and taxonomic implications. *Plant Systematics and Evolution*, 298(7), 1371–1381.
- Boshier, D., Dompreh, D., & Swaine, M. (Eds.). (2011). *Talbotiella gentii*: Genetic variation and conservation. A case study and teacher's notes. In *Forest Genetic Resources Training Guide*. Bioversity International.
- Brackley, F. E. (2006). *Jesup's milk-vetch* (Astragalus robbinsii *var.* jesupi): *Recovery plan* [US Fish and Wildlife Service]. New Hampshire Natural Heritage Inventory.
- Bradgø, M. (1957). Interspecific crosses in *Lupinus* cytology and inheritance of flower colour. *Hereditas*, 43(2), 338–356.
- Brewbaker, J. L. (1954). Incompatibility in autotetraploid *Trifolium repens* L. I. Competition and self-compatibility. *Genetics*, *39*(3), 307.
- Brewbaker, J. L. (1955). Studies of oppositional allelism in *Trifolium nigrescens*. *Hereditas*, 41(3-4), 367–375.
- Brown, J., Enright, N., & Miller, B. (2003). Seed production and germination in two rare and three common co-occurring *Acacia* species from south-east Australia. *Austral Ecology*, *28*(3), 271–280.
- Bruneau, A., & Anderson, G. J. (1988). Reproductive biology of diploid and triploid *Apios americana* (Leguminosae). *American Journal of Botany*, *75*(12), 1876–1883.
- Brunsberg, K., & others. (1977). Biosystematics of the Lathyrus pratensis complex.
- Bubar, J. (1958). An association between variability in ovule development within ovaries and selfincompatibility in *Lotus* (Leguminosae). *Canadian Journal of Botany*, *36*(1), 65–72.
- Bulgannawar, G. N., Math, B. B. M., & others. (1991). Advances in tropical *Acacia* research. In J. W. Turnbull (Ed.), *ACIAR Proceedings* (pp. 110–115).
- Bullock, S. H. (1985). Breeding systems in the flora of a tropical deciduous forest in Mexico. *Biotropica*, *17*(4), 287–301.
- Bumrungsri, S., Harbit, A., Benzie, C., Carmouche, K., Sridith, K., & Racey, P. (2008). The pollination ecology of two species of *Parkia* (Mimosaceae) in southern Thailand. *Journal of Tropical Ecology*, 24(5), 467–475.
- Burt, R., Williams, W., & Compton, J. (1973). Variation within naturally occurring Townsville stylo (*Stylosanthes humilis*) population; changes in population structure and some agronomic implications. *Australian Journal of Agricultural Research*, 24(5), 703–713.

- Bystricky, M., Schultze-Kraft, R., & Peters, M. (2010). Studies on the pollination biology of the tropical forage legume shrub *Cratylia argentea*. *Tropical Grasslands*, 44, 246–252.
- Byth, D. (1964). Breeding system and chromosome number in *Lotononis bainesii* Baker. *Nature, 202*(4934), 830–830.
- Cane, J. H. (2006). An evaluation of pollination mechanisms for purple prairie-clover, *Dalea purpurea* (Fabaceae: Amorpheae). *The American Midland Naturalist*, *156*(1), 193–197.
- Cane, J. H., Weber, M., & Miller, S. (2012). Breeding biologies, pollinators, and seed beetles of two prairie-clovers, *Dalea ornata* and *Dalea searlsiae* (Fabaceae: Amorpheae), from the Intermountain West, USA. Western North American Naturalist, 72(1), 16–20.
- Cardel, Y. J., & Koptur, S. (2010). Effects of florivory on the pollination of flowers: An experimental field study with a perennial plant. *International Journal of Plant Sciences*, *171*(3), 283–292.
- Carvalho, D. A., & Oliveira, P. E. (2003). Reproductive biology and pollination of *Senna sylvestris* (Vell.) HS Irwin & Barneby (Leguminosae, Caesalpinoideae). *Brazilian Journal of Botany*, *26*(3), 319–328.
- Casey, N. M., Milbourne, D., Barth, S., Febrer, M., Jenkins, G., Abberton, M. T., Jones, C., & Thorogood, D. (2010). The genetic location of the self-incompatibility locus in white clover (*Trifolium repens* L.). *Theoretical and Applied Genetics*, *121*(3), 567–576.
- Chamberlain, J. R. (1998). *Calliandra calothyrsus*: Reproductive biology in relation to its role as an important multipurpose tree. In *Reproductive Biology in Systematics, Conservation and Economic Botany* (pp. 439–448).
- Chee, W. C., & Devendra, C. (1983). Research on *Leucaena* forage production in Malaysia. *Leucaena Research in the Asian Pacific Region*.
- Chen, G. Q., Crawford, D., Huang, H., & GE, X. J. (2009). Genetic structure and mating system of *Ammopiptanthus mongolicus* (Leguminosae), an endangered shrub in north-western China. *Plant Species Biology*, 24(3), 179–188.
- Chen, J., & Gilbert, M. (2006). *Flora of China* (Vol. 10). Missouri Botanical Garden. http://efloras.org/
- Choudhury, B. I., Khan, M. L., & Dayanandan, S. (2014). Functional androdioecy in critically endangered *Gymnocladus assamicus* (Leguminosae) in the eastern Himalayan region of northeast India. *PloS One*, 9(2), e87287.
- Chow, K., & Crowder, L. (1973). Hybridization of *Desmodium* species. *Euphytica*, 22(2), 399-404.
- Clarke, H. D., Seigler, D. S., & Ebinger, J. E. (1990). *Acacia constricta* (Fabaceae: Mimosoideae) and related species from the southwestern US and Mexico. *American Journal of Botany*, 77(3), 305–315.
- Clements, R., & Ludlow, M. (1977). Frost avoidance and frost resistance in *Centrosema virginianum*. *Journal of Applied Ecology*, 14(2), 551–566.
- Clewell, A. F. (1966). Native north american species of *Lespedeza* (Leguminosae). *Rhodora*, 68(775), 359–405.
- Coaldrake, J. (1971). Variation in some floral, seed, and growth characteristics of *Acacia harpophylla* (Brigalow). *Australian Journal of Botany*, *19*(3), 335–352.
- Coe, H. S., & Martin, J. (1920). *Sweet-clover seed* [Bulletin of the United States Department of Agriculture]. USDA.
- Cole, C. T., & Biesboer, D. D. (1992). Monomorphism, reduced gene flow, and cleistogamy in rare and common species of *Lespedeza* (Fabaceae). *American Journal of Botany*, *79*(5), 567–575.

- Correia, M. C. L. (2012). *Reproductive Biology of Australian* Acacias *in Portugal* [PhD Thesis]. Universidade de Coimbra.
- Costa, C. B., Costa, J. A., de Queiroz, L. P., & Borba, E. L. (2013). Self-compatible sympatric *Chamae-crista* (Leguminosae-Caesalpinioideae) species present different interspecific isolation mechanisms depending on their phylogenetic proximity. *Plant Systematics and Evolution*, 299(4), 699–711.
- Costa, C. B. N., Lambert, S. M., Borba, E. L., & De Queiroz, L. P. (2007). Post-zygotic reproductive isolation between sympatric taxa in the *Chamaecrista desvauxii* complex (Leguminosae-Caesalpinioideae). *Annals of Botany*, 99(4), 625–635.
- Couderc, H. (1971). Etude expérimentale de la reproduction de l'Anthyllis vulneraria L. Bulletin de La Société Botanique de France, 118(5-6), 359-373.
- Cram, W. H. (1952). Parent-seedling characteristics and relationships in *Caragana arborescens* Lam. *Scientific Agriculture*, *32*(7), 380–402.
- Cram, W. H. (1955). Self-compatibility of *Caragana arborescens* Lam. *Canadian Journal of Botany*, 33(2), 149–155.
- Culjak, M. V., & Forrest, J. (2018). Specialist vs. Generalist: Which is the best pollinator? The reproductive biology and pollination ecology of the Nevada peavine (Lathyrus lanszwertii).
- Culley, T. M., & Klooster, M. R. (2007). The cleistogamous breeding system: A review of its frequency, evolution, and ecology in angiosperms. *The Botanical Review*, 73(1), 1.
- Dahlgren, R. (1963). Studies on Aspalathus: Phytogeographical aspects. Carl Bloms Boktryckeri.
- Darwin, C. (1877). The effects of cross and self fertilisation in the vegetable kingdom. D. Appleton.
- da Silva, A. L. G., Chaves, S. R., & Brito, J. M. (2011). Reproductive biology of *Bowdichia virgilioides* Kunth (Fabaceae). *Acta Scientiarum. Biological Sciences*, 33(4), 463-470.
- de Artucio, I. (2017). Food and Agriculture Organization of the United States. In Adesmia bicolor (Poiret) De Candolle. http://www.fao.org/ag/agp/AGPC/doc/gbase/data/pf000509.html
- Decker, K., & Anderson, D. G. (2004). Astragalus anisus *M.E. Jones (Gunnison milkvetch): A Technical Conservation Assessment* [USDA Forest Service]. Colorado Natural Heritage Program.
- de las Heras, M., Hidalgo, P., & Ubera, J. (2001). Stigmatic cuticle in *Hedysarum glomeratum*: Structure and function. *International Journal of Developmental Biology*, *45*(S1), S41–S42.
- Delgado-Salinas, A. (2000). New species of Mexican *Phaseolus* (Fabaceae). *Systematic Botany*, *25*(3), 414–436.
- de Lima Leite, A. V., & Machado, I. C. (2010). Reproductive biology of woody species in Caatinga, a dry forest of northeastern Brazil. *Journal of Arid Environments*, 74(11), 1374–1380.
- de Oliveira Barros, E. C. (2013). Limitação de polinizadores e mecanismo de autoincompatibilidade de ação tardia como causas da baixa formação de frutos em duas espécies simpátricas de *Inga* (FabaceaeMimosoideae) na Amazônia Central. *Rodriguésia-Instituto de Pesquisas Jardim Botânico Do Rio de Janeiro*, 64(1), 37–47.
- de Oliveira Santos-Garcia, M., Resende, R. M. S., Chiari, L., Zucchi, M. I., & de Souza, A. P. (2011). Mating systems in tropical forages: *Stylosanthes capitata* Vog. And *Stylosanthes guianensis* (Aubl.) Sw. *Euphytica*, 178(2), 185–193.
- de Queiroz, L. P., da SILVA, M. M., Ramos, A. K., & Pizarro, E. A. (1997). Estudos reprodutivos em *Cratylia argentea* (desv.) o. Kuntze e *Cratylia mollis* mart. Ex Benth. (Leguminosae-Papilionoideae). *Pasturas Tropicales*, 19, 20-citation\_lastpage.

- Diallo, A. M., Nielsen, L. R., Hansen, J. K., Ræbild, A., & Kjær, E. D. (2015). Study of quantitative genetics of gum arabic production complicated by variability in ploidy level of *Acacia senegal*. *Tree Genetics & Genomes*, *11*(4), 1–13.
- Diallo, B. O., Mckey, D., Chevallier, M., Joly, H. I., & Hossaert-Mckey, M. (2008). Breeding system and pollination biology of the semidomesticated fruit tree, *Tamarindus indica* L. (Leguminosae: Caesalpinioideae): Implications for fruit production, selective breeding, and conservation of genetic resources. *African Journal of Biotechnology*, 7(22), 4068–4075.
- Domínguez, V. E., & Gibbs, P. E. (1975). Estudio experimental de la reproducción en las especies anuales europeas del género *Hippocrepis* L. *Lagascalia*, *5*(1), 35–46.
- Doungyotha, Y., & Owens, J. N. (2002). The reproductive biology and reproductive success of *Ptero-carpus macrocarpus* Kurz. *Biotropica*, *34*(1), 58–67.
- Drummond, C. S., & Hamilton, M. B. (2007). Hierarchical components of genetic variation at a species boundary: Population structure in two sympatric varieties of *Lupinus microcarpus* (Leguminosae). *Molecular Ecology*, *16*(4), 753–769.
- Duminil, J., Danou, K., Kaviriri, D. K., Gillet, P., Loo, J., Doucet, J.-L., & Hardy, O. (2016). Relationships between population density, fine-scale genetic structure, mating system and pollen dispersal in a timber tree from African rainforests. *Heredity*, 116(3), 295–303.
- Dunn, D. B. (1971). A case of long range dispersal and "rapid speciation" in *Lupinus*. *Transactions of the Missouri Academy of Science*, *5*, 26–38.
- Eade, R., McDonald, F., & Simes, J. (1973). Extractives of Australian timbers. XIV. Triterpene glycosides of Acacia myrtifolia. Australian Journal of Chemistry, 26(4), 839–844.
- East, E. M. (1940). The distribution of self-sterility in the flowering plants. *Proceedings of the American Philosophical Society*, *82*(4), 449–518.
- Edwardson, J. (1967). Cytoplasmic male sterility and fertility restoration in *Crotalaria mucronata*. *Journal of Heredity*, *58*(5), 266–268.
- Elias, T. S. (1974). The genera of Mimosoideae (Leguminosae) in the southeastern United States. *Journal of the Arnold Arboretum*, 55(1), 67–118.
- Eliášová, A., Trávníček, P., Mandák, B., & Münzbergová, Z. (2014). Autotetraploids of *Vicia cracca* show a higher allelic richness in natural populations and a higher seed set after artificial selfing than diploids. *Annals of Botany*, *113*(1), 159–170.
- Elliot, M., Harding, J., & Mankinen, C. (1974). Genetics of *Lupinus* VI. Reproductive barriers in the Nanus group. *Taxon*, 23(4), 585–594.
- Encyclopedia of Life. (n.d.). In Desert milkvetch. https://eol.org/pages/639185
- Eshel, Y. (1968). Flower development and pollen viability of chickpea (*Cicer arietinum* L.). *The Isreal Journal of Agricultural Research*, *18*(1), 31–33.
- Etcheverry, A., & Aleman, C. (2005). Reproductive biology of *Erythrina falcata* (Fabaceae : Papilionoideae). *Biotropica*, *37*(1), 54-63.
- Etcheverry, A. V., Alemán, M. M., & Fleming, T. F. (2008). Flower morphology, pollination biology and mating system of the complex flower of *Vigna caracalla* (Fabaceae: Papilionoideae). *Annals of Botany*, 102(3), 305–316.
- Eynard, C., & Galetto, L. (2002). Pollination ecology of *Geoffroea decorticans* (Fabaceae) in central Argentine dry forest. *Journal of Arid Environments*, *51*(1), 79–88.

- Faivre, A. (2013). Plant Reproductive Biology of Chasmogamous and Cleistogamous Flowers of Federally Threatened Clitoria fragrans (Fabaceae) Along the Lake Wales Ridge in Florida. (Abstract only) [Poster].
- Feinsinger, P., Linhart, Y. B., Swarm, L. A., & Wolfe, J. A. (1979). Aspects of the pollination biology of three *Erythrina* species on Trinidad and Tobago. *Annals of the Missouri Botanical Garden*, 451–471.
- Fenster, C. B. (1995). Mirror image flowers and their effect on outcrossing rate in *Chamaecrista fascic-ulata* (Leguminosae). *American Journal of Botany*, *82*(1), 46–50.
- Ferreira, N. R., & Franke, L. B. (2011). Estudos morfo-anatômicos relacionados à dormência em sementes de *Adesmia tristis* Vogel (Fabaceae). *Revista Brasileira de Sementes*, 33(3), 447–453.
- Frankie, G. W., Opler, P. A., & Bawa, K. S. (1976). Foraging behaviour of solitary bees: Implications for outcrossing of a neotropical forest tree species. *The Journal of Ecology*, 1049–1057.
- Free, J. B., & others. (1970). Insect pollination of crops. Academic press.
- Fryxell, P. A. (1957). Mode of reproduction of higher plants. The Botanical Review, 23(3), 135–233.
- Fukui, J., & Takahashi, M. (1975). Ecological differentiation in the flowering and fruiting habits of *Amphicarpaea edgeworthii* in Japan. *Journal of the Faculty of Agriculture-Iwate University (Japan)*.
- Galetto, L., Bernardello, G., Isele, I. C., Vesprini, J., Speroni, G., & Berduc, A. (2000). Reproductive biology of *Erythrina crista-galli* (Fabaceae). *Annals of the Missouri Botanical Garden*, 87(2), 127– 145.
- Gallardo, R., Dominguez, E., & Munoz, J. M. (1993). The heterochronic origin of the cleistogamous flower in *Astragalus cymbicarpos* (Fabaceae). *American Journal of Botany*, *80*(7), 814–823.
- Gallardo, R., Dominguez, E., & Munoz, J. M. (1994). Pollen-ovule ratio, pollen size, and breeding system in Astragalus (Fabaceae) subgenus Epiglottis: A pollen and seed allocation approach. American Journal of Botany, 81(12), 1611–1619.
- Galloni, M., Podda, L., Vivarelli, D., & Cristofolini, G. (2007). Pollen presentation, pollen-ovule ratios, and other reproductive traits in Mediterranean Legumes (Fam. Fabaceae-Subfam. Faboideae). *Plant Systematics and Evolution*, 266(3), 147–164.
- Gasmanova, N., Lebeda, A., Doležalová, I., Cohen, T., Pavlíček, T., Fahima, T., & Nevo, E. (2007). Genome size variation of *Lotus peregrinus* at evolution canyon I microsite, lower Nahal Oren, Mt. Carmel, Israel. *Acta Biologica Cracoviensia Series Botanica*, 49(1), 39–46.
- Gassama-Dia, Y., Sane, D., & N'Doye, M. (2003). Reproductive biology of *Faidherbia albida* (Del.) A. Chev. *Silva Fennica*, *37*(4), 429–436.
- Geer, S., Tepedino, V. J., Griswold, T. L., & Bowlin, W. (1995). Pollinator sharing by three sympatric milkvetches, including the endangered species *Astragalus montii*. *The Great Basin Naturalist*, 19– 28.
- George, N., Byrne, M., & Yan, G. (2008). Mixed mating with preferential outcrossing in *Acacia saligna* (Labill.) H. Wendl. (Leguminosae: Mimosoideae). *Silvae Genetica*, *57*(3), 139.
- Gélvez-Zúñiga, I., Neves, A. C., Teixido, A. L., & Fernandes, G. W. (2018). Reproductive biology and floral visitors of *Collaea cipoensis* (Fabaceae), an endemic shrub of the rupestrian grasslands. *Flora-Morphology, Distribution, Functional Ecology of Plants, 238*, 129–137.
- Gibbs, P. E., & Bianchi, M. B. (1999). Does late-acting self-incompatibility (LSI) show family clustering? Two more species of Bignoniaceae with LSI: *Dolichandra cynanchoides* and *Tabebuia nodosa*. *Annals of Botany*, 84(4), 449–457.

- Gibbs, P., & Sassaki, R. (1998). Reproductive biology of *Dalbergia miscolobium* Benth. (Leguminosae-Papilionoideae) in SE Brazil: The effects of pistillate sorting on fruit-set. *Annals of Botany*, *81*(6), 735–740.
- Giles, W. L. (1949). *The morphological aspects of self-sterility in* Lotus corniculatus *L.* [PhD thesis]. University of Missouri.
- Godt, M. J. W., & Hamrick, J. (1997). Estimation of mating system parameter of *Albizia julibrissin* (Fabaceae). *International Journal of Forest Genetics*, 4(4), 217–222.
- Golubov, J., Eguiarte, L. E., Mandujano, M. C., López-Portillo, J., & Montaña, C. (1999). Why be a honeyless honey mesquite? Reproduction and mating system of nectarful and nectarless individuals. *American Journal of Botany*, *86*(7), 955–963.
- Gomes da Silva, A. L., Ramos Chaves, S., & Mesquita Brito, J. (2011). Reproductive biology of *Bowdichia virgilioides* Kunth (Fabaceae). *Acta Scientiarum Biological Sciences*, 33(4), 463–470.
- Gomez, S. M., & Kalamani, A. (2003). Butterfly pea (*Clitoria ternatea*): A nutritive multipurpose forage legume for the tropicsan overview. *Pakistan Journal of Nutrition*, *2*(6), 374–379.
- Gottsberger, G., Camargo, J. M., & Silberbauer-Gottsberger, I. (1988). A bee pollinated tropical community: The beach dune vegetation of Ilha de São Luís, Maranhão, Brazil. *Botanische Jahrbücher Für Systematik, Pflanzengeschichte, Und Pflanzengeographie, 109*(4), 469–500.
- Gómez, P. P., Lillo, D., & González, A. V. (2012). Pollination and breeding system in *Adesmia bijuga* Phil. (Fabaceae), a critically endangered species in Central Chile. *Gayana Botanica*, 69(2), 286.
- Grant, P., Grant, K. T., & Grant, B. R. (1991). *Erythrina velutina* and the colonization of remote islands. *Noticias de Galapagos*, *50*, 3–5.
- Green, T. W., & Bohart, G. E. (1975). The pollination ecology of Astragalus cibarius and Astragalus utahensis (Leguminosae). American Journal of Botany, 62(4), 379–386.
- Gross, C. (1990). The breeding systems of three cooccurring legumes: *Dillwynia hispida, Dillwynia uncinata*, and *Pultenaea densifolia* (Leguminosae, Papilionoideae). *Australian Journal of Botany*, 38(2), 207–215.
- Gross, C. (1992). Floral traits and pollinator constancy: Foraging by native bees among three sympatric legumes. *Australian Journal of Ecology*, *17*(1), 67–74.
- Gross, C. (1993). The reproductive ecology of *Canavalia rosea* (Fabaceae) on Anak Krakatau, Indonesia. *Australian Journal of Botany*, *41*(5), 591–599.
- Gunn, C. R., & Kluve, J. (1976). Androecium and pistil characters for tribe Vicieae (Fabaceae). *Taxon*, *25*(5-6), 563–575.
- Haddock, R. C., & Chaplin, S. J. (1982). Pollination and seed production in two phenologically divergent prairie legumes (*Baptisia leucophaea* and *B. leucantha*). *American Midland Naturalist*, 175–186.
- Haines, A. (2001). Oxytropis campestris (*L.*) *DC. Var.* johannensis *Fern. St. John River Oxytrope*. New England Wild Flower Society.
- Hall, H. G., & Avila, L. (2016). *Megachile sculpturalis*, the giant resin bee, overcomes the blossom structure of Sunn Hemp (*Crotalaria juncea*) that impedes pollination. *Journal of Melittology*, 65, 1–11.
- Hanson, C. H. (1953). *Lespedeza stipulacea*: Stamen morphology, meiosis, microgametogenesis, and fertilization 1. *Agronomy Journal*, *45*(5), 200–203.
- Hanson, C. H., & Cope, W. A. (1955). Reproduction in the cleistogamous flowers of ten perennial species of *Lespedeza*. *American Journal of Botany*, *42*(7), 624–627.

- Harding, J., Mankinen, C., & Elliott, M. (1974). Genetics of *Lupinus* VII. Outcrossing, autofertility, and variability in natural populations of the Nanus group. *Taxon*, *23*(5-6), 729–738.
- Hartley, M. L., Tshamekeng, E., & Thomas, S. M. (2002). Functional heterostyly in *Tylosema esculentum* (Caesalpinioideae). *Annals of Botany*, *89*(1), 67–76.
- Hawkins, C., & Evans, A. (1973). Elucidating the behaviour of pollen tubes in intra-and interspecific pollinations of *Phaseolus vulgaris* L. And *P. coccineus* Lam. *Euphytica*, *22*(2), 378–385.
- Heenan, P. B., & De Lange, P. J. (1999). Reproductive biology, ecology and conservation of *Carmichaelia williamsii* (Fabaceae), a vulnerable legume from New Zealand. *Pacific Conservation Biology*, 5(3), 179–190.
- Heering, J. H. (1994). The reproductive biology of three perennial *Sesbania* species (Leguminosae). *Euphytica*, *74*(1-2), 143–148.
- Heinrichs, D. (1963). Creeping alfalfas. In Advances in agronomy (Vol. 15, pp. 317–337). Elsevier.
- Heithaus, E. R., Opler, P. A., & Baker, H. G. (1974). Bat activity and pollination of *Bauhinia pauletia*: Plant-pollinator coevolution. *Ecology*, *55*(2), 412–419.
- Helenurm, K., & Schaal, B. A. (1996). Genetic and maternal effects on offspring fitness in *Lupinus texensis* (Fabaceae). *American Journal of Botany*, 83(12), 1596–1608.
- Herendeen, P. S., Lewis, G. P., & Bruneau, A. (2003). Floral morphology in Caesalpinioid legumes: Testing the monophyly of the "Umtiza clade". International Journal of Plant Sciences, 164(S5), S393– S407.
- Hernandez, K. L. R. (2012). *Biología floral dela ecotipo yapacaní de* Cratylia argentea (*Fabaceae, Faboideae*) [PhD Thesis]. Universidad del Valle.
- Hernández, H. M. (1982). Female sterility in *Erythrina montana*. Allertonia, 3, 71–76.
- Hernández, H. M., & Toledo, V. M. (1979). The role of nectar robbers and pollinators in the reproduction of *Erythrina leptorhiza*. *Annals of the Missouri Botanical Garden*, 66(3), 512–520.
- Heyn, C. C., & Raviv, V. (1966). Experimental taxonomic studies in the genus *Scorpiurus* (Papilionaceae). *Bulletin of the Torrey Botanical Club*, 93(4), 259–267.
- Hoc, P. S., Drewes, S. I., & Amela García, M. T. (2003). Biología floral, sistema reproductivo y éxito reproductivo de *Macroptilium fraternum* (Fabaceae). *Revista de Biología Tropical*, *51*(2), 369–380.
- Hokche, O., & Ramirez, Y. N. (2016). Autoincompatibilidad en especies de Bauhinia L. (Fabaceae: Caesalpiniodeae) en Venezuela. 8, 105-108.
- Horovitz, A., Meiri, L., & Beiles, A. (1976). Effects of ovule positions in Fabaceous flowers on seed set and outcrossing rates. *Botanical Gazette*, *137*(3), 250–254.
- Howard, A., Howard, G. L., & Abdur, R. (1919). Studies in the pollination of Indian crops. *Memoirs of the Department of Agriculture in India. Botanical Series.*, 10, 195–220.
- Hufford, K. M., & Hamrick, J. (2003). Viability selection at three early life stages of the tropical tree, *Platypodium elegans* (Fabaceae, Papilionoideae). *Evolution*, *57*(3), 518–526.
- Hutton, E. (1960). Flowering and pollination in *Indigofera spicata, Phaseolus lathyroides, Desmodium uncinatum,* and some other tropical pasture legumes. *Journal of Experimental Agriculture, 28*(111), 235–243.
- Hutton, E. (1970). Tropical pastures. Advances in Agronomy, 22, 1-73.
- Hutton, E. M., & Eddie, T. (1982). *Leucaena esculenta* and *L. trichodes* some similarities and differences. *Leucaena Research Reports*, *3*, 12–13.

- Imrie, B., & Hutton, E. (1978). The inheritance of flower color in *Macroptilium atropurpureum*. *Journal* of Heredity, 69(1), 54–56.
- Jacobson, A. L. (2005). Arthur Lee Jacobson. In Plant of the Month: July 2005, Wattle. http:// www.arthurleej.com/p-o-m-July05.html
- Jaimes, I., & Ramirez, N. (1999). Breeding systems in a secondary deciduous forest in Venezuela: The importance of life form, habitat, and pollination specificity. *Plant Systematics and Evolution*, 215(1), 23–36.
- Janzen, D. H. (1974). Swollen-thorn *Acacias* of Central America. *Smithsonian Contributions to Botany*, 13.
- Janzen, D. H. (1977). Variation in seed weight in Costa Rican Cassia grandis (Leguminosae). Variación en el peso de la semilla costarricense de Cassia grandis (Leguminosae). Tropical Ecology, 18(2), 177– 186.
- Janzen, D. H. (1978). Inter-and intra-crop variation in seed weight of Costa Rican *Ateleia herbert-smithii* Pitt. (Leguminosae). *Brenesia*, *14*(15), 311–323.
- Karoly, K. (1992). Pollinator limitation in the facultatively autogamous annual, *Lupinus nanus* (Leguminosae). *American Journal of Botany*, 79(1), 49–56.
- Karron, J. D. (1989). Breeding systems and levels of inbreeding depression in geographically restricted and widespread species of *Astragalus* (Fabaceae). *American Journal of Botany*, *76*(3), 331–340.
- Kaul, V., Koul, A., & Sharma, M. (2000). The underground flower. Current Science, 78(1), 39-44.
- Kaye, T. N. (1999). From flowering to dispersal: Reproductive ecology of an endemic plant, Astragalus australis var. olympicus (Fabaceae). American Journal of Botany, 86(9), 1248–1256.
- Kempf, K., Grieder, C., Walter, A., Widmer, F., Reinhard, S., & Kölliker, R. (2015). Evidence and consequences of self-fertilisation in the predominantly outbreeding forage legume *Onobrychis viciifolia*. *BMC Genetics*, 16(1), 117.
- Kenrick, J., & Knox, R. B. (1989). Quantitative analysis of self-incompatibility in trees of seven species of *Acacia*. *Journal of Heredity*, *80*(3), 240–245.
- Keys, R. N. (1993). *Mating systems and pollination biology of velvet mesquite* (Prosopis velutina *Wooten*) [PhD thesis]. University of Arizona.
- Keys, R. N., & Smith, S. E. (1994). Mating system parameters and population genetic structure in pioneer populations of *Prosopis velutina* (Leguminosae). *American Journal of Botany*, *81*(8), 1013–1020.
- Kiill, L. H. P., & Drumond, M. A. (2001). Biologia floral e sistema reprodutivo de *Gliricidia sepium* (Jacq.) Steud. (Fabaceae-Papilionoidae) na região de Petrolina, Pernambuco. *Ciência Rural*, 31(4), 597–601.
- Kirchner, D. (1905). Über die Wirkung der Selbstbestaubung bei den Papilionaceen. Ulmer.
- Kirk, L. E. (1925). Self-pollination of sweet clover. *Scientific Agriculture*, 6(4), 109–112.
- Kita, F. (1965). Studies on the genus *Melilotus* (sweetclover) with special reference to interrelationships among species from a cytological point of view. *Journal of the Faculty of Agriculture, Hokkaido* University, 54(2), 23–122.
- Kittelson, P. M., & Maron, J. L. (2000). Outcrossing rate and inbreeding depression in the perennial yellow bush lupine, *Lupinus arboreus* (Fabaceae). *American Journal of Botany*, *87*(5), 652–660.
- Kodela, P. (2000). Flora of New South Wales (rev., Vol. 2). NSW University Press.

- Kokkoris, I., & Arianoutsou, M. (2000). Flowering and fruiting phenology of four herbaceous species of Leguminosae in a burned *Pinus halepensis* forest of Attica, Greece. *Journal of Mediterranean Ecology*, 1, 193–200.
- Koptur, S. (1984). Outcrossing and pollinator limitation of fruit set: Breeding systems of neotropical *Inga* trees (Fabaceae, Mimosoideae). *Evolution*, *38*(5), 1130–1143.
- Kozuharova, E. K. (1999). On the reproductive biology of *Onobrychis pindicola* Hausskn. Subsp. *urumovii* Deg. & Dren. (Fabaceae). *Flora-Morphology, Distribution, Functional Ecology of Plants, 9, 291.*
- Kudo, G., & Harder, L. (2005). Floral and inflorescence effects on variation in pollen removal and seed production among six legume species. *Functional Ecology*, *19*(2), 245–254.
- Kudo, G., & Molau, U. (1999). Variations in reproductive traits at inflorescence and flower levels of an arctic legume, *Astragalus alpinus* L.: Comparisons between a subalpine and an alpine population. *Plant Species Biology*, 14(3), 181–191.
- Kupicha, F. K. (1977). The delimitation of the tribe Vicieae (Leguminosae) and the relationships of *Cicer* L. *Botanical Journal of the Linnean Society*, 74(2), 131–162.
- Ladizinsky, G., & Adler, A. (1976). Genetic relationships among the annual species of *Cicer* L. *Theoretical* and *Applied Genetics*, 48(4), 197–203.
- Lambeth, V. N. (1950). Some factors influencing pod set and yield of the lima bean.
- Lancaster, R. R. (1949). Clovers for texas pastures. *Bulletin/Texas Agricultural Extension Service; No.* 168.
- Laporta, C. (2005). Floral biology and reproductive system of enantiostylous *Senna corymbosa* (Caesalpiniaceae). *Revista de Biología Tropical*, *53*(1-2), 49–61.
- Lassen, K. M., Kjær, E. D., Ouédraogo, M., Dupont, Y. L., & Nielsen, L. R. (2018). Controlled pollinations reveal self-incompatibility and inbreeding depression in the nutritionally important parkland tree, *Parkia biglobosa*, in Burkina Faso. *Journal of Pollination Ecology*, 24(16), 144–156.
- Latouche-Hallé, C., Ramboer, A., Bandou, E., Caron, H., & Kremer, A. (2004). Long-distance pollen flow and tolerance to selfing in a neotropical tree species. *Molecular Ecology*, *13*(5), 1055–1064.
- Latting, J. (1961). The biology of *Desmanthus illinoensis*. *Ecology*, 42(3), 487–493.
- Lau, C. P. Y., Saunders, R. M. K., & Ramsden, L. (2009). Floral biology, breeding systems and population genetic structure of three climbing *Bauhinia* species (Leguminosae: Caesalpinioideae) in Hong Kong, China. *Journal of Tropical Ecology*, 25(2), 147–159.
- Leduc, N., Douglas, G. C., Monnier, M., & Connolly, V. (1990). Pollination in vitro effects on the growth of pollen tubes, seed set, and gametophytic self-incompatibility in *Trifolium pratense* L. And *Trifolium repens* L. *Theoretical and Applied Genetics*, *80*(5), 657–664.
- Leela, N. K., & Shafeekh, K. M. (2008). Fenugreek. In *Chemistry of Spices* (pp. 242–258). CAB International.
- Leite, A. C. R., Araújo, T. G., Carvalho, B. de M., Maia, M. B. S., & Lima, V. L. de M. (2010). Characterization of the antidiabetic role of *Parkinsonia aculeata* (Caesalpineaceae). *Evidence-Based Complementary and Alternative Medicine*, 2011.
- Lemus-Jiménez, L. J., & Ramírez, N. (2005). Sistemas reproductivos de las plantas en tres hábitats de la planicie costera de Paraguaná, Venezuela. *Revista de Biología Tropical, 53*(3-4), 415–430.
- Lewis, G., & Gibbs, P. (1999). Reproductive biology of *Caesalpinia calycina* and *C. pluviosa* (Leguminosae) of the Caatinga of north-eastern Brazil. *Plant Systematics and Evolution*, *217*(1), 43–53.
- Liengsiri, C., Boyle, T. J., & Yeh, F. C. (1998). Mating system in *Pterocarpus macrocarpus* Kurz in Thailand. *Journal of Heredity*, 89(3), 216–221.
- Linares, L. J., & Koptur, S. (2010). Floral biology and breeding system of the crenulate leadplant, Amorpha herbacea var. crenulata, an endangered South Florida pine rockland endemic. Natural Areas Journal, 30(2), 138–147.
- Liu, H., & Koptur, S. (2003). Breeding system and pollination of a narrowly endemic herb of the Lower Florida Keys: Impacts of the urban-wildland interface. *American Journal of Botany*, *90*(8), 1180– 1187.
- Lloyd, D., & Schoen, D. (1992). Self-fertilization and cross-fertilization in plants. I. Functional dimensions. *International Journal of Plant Sciences*, *153*(3), 358–369.
- Lord, E., & Heslop-Harrison, Y. (1984). Pollen-stigma interaction in the Leguminosae: Stigma organization and the breeding system in *Vicia faba* L. *Annals of Botany*, *54*(6), 827–836.
- Lord, E. M. (1981). Cleistogamy: A tool for the study of floral morphogenesis, function and evolution. *The Botanical Review*, *47*(4), 421–449.
- Loveless, M., Hamrick, J., & Foster, R. (1998). Population structure and mating system in *Tachigali* versicolor, a monocarpic neotropical tree. *Heredity*, 81(2), 134–143.
- López, J., Rodríguez-Riaño, T., Ortega-Olivencia, A., Antonio Devesa, J., & Ruiz, T. (1999). Pollination mechanisms and pollen-ovule ratios in some Genisteae (Fabaceae) from Southwestern Europe. *Plant Systematics and Evolution*, 216(1), 23–47.
- Lukefahr, M., & Fryxell, P. A. (1967). Content of gossypol in plants belonging to genera related to cotton. *Economic Botany*, 21(2), 128–131.
- Lundqvist, A. (1993). The self-incompatibility system in *Lotus tenuis* (Fabaceae). *Hereditas*, 119(1), 59–66.
- Lydaki, M. E., & Vlahos, J. C. (1998). Pollination Mechanisms of Ebenus cretica [Poster].
- Ma, X., Milne, R., Zhou, H., Fang, J., & Zha, H. (2017). Floral nectar of the obligate outcrossing *Canavalia gladiata*, (Jacq.) DC. (Fabaceae) contains only one predominant protein, a class III acidic chitinase. *Plant Biology*, 19(5), 749–759.
- Machado, I. C., Lopes, A. V., & Sazima, M. (2006). Plant sexual systems and a review of the breeding system studies in the Caatinga, a Brazilian tropical dry forest. *Annals of Botany*, *97*(2), 277–287.
- Martin, M., & Lee, T. (1993). Self-pollination and resource availability affect ovule abortion in *Cassia fasciculata* (Caesalpiniaceae). *Oecologia*, *94*(4), 503–509.
- Martínez-Sánchez, J. J., Segura, F., Aguado, M., Franco, J. A., & Vicente, M. J. (2011). Life history and demographic features of Astragalus nitidiflorus, a critically endangered species. Flora-Morphology, Distribution, Functional Ecology of Plants, 206(5), 423–432.
- McKee, R., & Hyland, H. (1941). Apetalous and petaliferous flowers in *Lespedeza* 1. *Agronomy Journal*, 33(9), 811–815.
- Meadow jointvetch. (2000). In Florida National Areas Inventory. https://www.fnai.org/
- Meehan, T. (1886). On the fertilization of *Cassia marilandica*. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 314–318.
- Meglic, V., & Smith, R. (1992). Self-incompatibility and seed set in colchicine-, nitrous oxide-, and sexually derived tetraploid red clover. *Crop Science*, *32*(5), 1133–1137.

- Meinke, R. J., & Kaye, T. N. (1992). Taxonomic assessment of *Astragalus tegetarioides* (Fabaceae) and a new related species from northern California. *Madroño*, *39*(3), 193–204.
- Mendel, J. G. (1886). Versuche über plflanzenhybriden verhandlungen des naturforschenden vereines in brünn, bd. IV für das jahr. *Journal of the Royal Horticultural Society*, *26*, 1–32.
- Miles, J. (1985). Evaluation of potential genetic marker traits and estimation of outcrossing rate in *Stylosanthes guianensis. Crop and Pasture Science*, *36*(2), 259–265.
- Millar, M. A., Coates, D. J., & Byrne, M. (2014). Extensive long-distance pollen dispersal and highly outcrossed mating in historically small and disjunct populations of *Acacia woodmaniorum* (Fabaceae), a rare banded iron formation endemic. *Annals of Botany*, 114(5), 961–971.
- Minahan, D. (2012). Does individual flower density and patch level conspecific neighbor density affect reproductive success in a "Golden Banner" Legume (Fabaceae)? [Master's Thesis]. University of Colorado at Boulder.
- Miri, R., & Bubar, J. (1966). Self-incompatibility as an outcrossing mechanism in birdsfoot trefoil (*Lotus corniculatus*). *Canadian Journal of Plant Science*, *46*(4), 411–418.
- Moffett, A. (1956). Genetical studies in *Acacias*. I. The estimation of natural crossing in black wattle. *Heredity*, *10*(1), 57–67.
- Morales, C., & Galetto, L. (2003). Influence of compatibility system and life form on plant reproductive success. *Plant Biology*, *5*(5), 567–573.
- Moran, G., Muona, O., & Bell, J. (1989). Breeding systems and genetic diversity in Acacia auriculiformis and A. crassicarpa. Biotropica, 21(3), 250–256.
- Morgan, A., Carthew, S. M., & Sedgley, M. (2002). Breeding system, reproductive efficiency and weed potential of *Acacia baileyana*. *Australian Journal of Botany*, *50*(3), 357–364.
- Morley, F. (1963). The mode of pollination in strawberry clover (*Trifolium fragiferum*). Animal Production Science, 3(8), 5–8.
- Morley, F. (1983). Validation of combinations in the Trifolium subterraneum complex (Fabaceae). Taxon.
- Morrison, D. A., & Myerscough, P. J. (1989). Temporal regulation of maternal investment in populations of the perennial legume *Acacia suaveolens*. *Ecology*, *70*(6), 1629–1638.
- Moura, D. C. (2008). *Interações entre plantas e abelhas nas matas ciliares do Rio São Francisco* [PhD Thesis]. Universidade Federal de Pernambuco.
- Moyle, L. C., Olson, M. S., & Tiffin, P. (2004). Patterns of reproductive isolation in three angiosperm genera. *Evolution*, *58*(6), 1195–1208.
- Munin, R. L., Teixeira, R. C., & Sigrist, M. R. (2008). Esfingofilia e sistema de reprodução de Bauhinia curvula Benth. (Leguminosae: Caesalpinioideae) em cerrado no Centro-Oeste brasileiro. 31(1), 15-25.
- Nair, R. M., Dundas, I. S., Wallwork, M., Verlin, D. C., Waterhouse, L., & Dowling, K. (2004). Breeding system in a population of *Trigonella balansae* (Leguminosae). *Annals of Botany*, 94(6), 883–888.
- Nair, R. M., Peck, D. M., Rowe, T. D., Dundas, I. S., & Dowling, K. (2007). Breeding system in *Trifolium glanduliferum* (Fabaceae). *New Zealand Journal of Agricultural Research*, *50*(4), 451–456.
- Nason, J., & Hamrick, J. (1997). Reproductive and genetic consequences of forest fragmentation: Two case studies of neotropical canopy trees. *Journal of Heredity*, *88*(4), 264–276.
- Navarro, L. (1999). Reproductive biology of *Anthyllis vulneraria subsp. vulgaris* (Fabaceae) in northwestern Iberian Peninsula. *Nordic Journal of Botany*, *19*(3), 281–287.

- Nayak, K. G., & Davidar, P. (2010). Pollinator limitation and the effect of breeding systems on plant reproduction in forest fragments. *Acta Oecologica*, *36*(2), 191–196.
- Neill, D. A. (1988). Experimental studies on species relationship in *Erythrina* (Leguminosae, Papilionoideae). *Annals of the Missouri Botanical Garden*, 75(3), 886–969.
- Neto, H., & Ferreira, P. (2013). Floral biology and breeding system of *Bauhinia forficata* (Leguminosae: Caesalpinioideae), a moth-pollinated tree in southeastern Brazil. *Brazilian Journal of Botany*, *36*(1), 55–64.
- Netshisaulu, T. (2012). *PlantZAfrica.Com*. https://www.plantzafrica.com/plantab/acaciatortilis .htm
- Newstrom, L., & Robertson, A. (2005). Progress in understanding pollination systems in New Zealand. New Zealand Journal of Botany, 43(1), 1–59.
- Nielsen, I., Fortune, H., & others. (1992). Flora Malesiana. Mimosaceae. (Vol. 11).
- Novosyelova, L. (1998). The bud cleistogamy of some annual species of the genus *Medicago* (Fabaceae). *Bot. Zhurn.(St. Petersburg)*, 83, 82–87.
- Oliveira, P., & Gibbs, P. (2000). Reproductive biology of woody plants in a cerrado community of Central Brazil. *Flora-Morphology, Distribution, Functional Ecology of Plants, 195*(4), 311–329.
- Oliveira, P., Gibbs, P., & Barbosa, A. (2004). Moth pollination of woody species in the Cerrados of Central Brazil: A case of so much owed to so few? *Plant Systematics and Evolution*, *245*(1-2), 41–54.
- O'Malley, D. M., & Bawa, K. S. (1987). Mating system of a tropical rain forest tree species. *American Journal of Botany*, *74*(8), 1143–1149.
- Orwa, C., Mutua, A., Kindt, R., Jamnadass, R., & Anthony, S. (2009). Agroforestree Database: A tree reference and selection guide version 4.0. http://www.worldagroforestry.org/sites/treedbs/ treedatabases.asp
- Parker, I. M., & Haubensak, K. A. (2002). Comparative pollinator limitation of two non-native shrubs: Do mutualisms influence invasions? *Oecologia*, *130*(2), 250–258.
- Parrish, J., & Bazzaz, F. (1979). Difference in pollination niche relationships in early and late successional plant communities. *Ecology*, *60*(3), 597–610.
- Pazy, B. (1984). Insect induced self-pollination. *Plant Systematics and Evolution*, 144(3), 315–320.
- Peng, C. I., & Chaw, S. M. (1986). *Alysicarpus rugosus* (Willd.) DC., A newly naturalized legume species in Taiwan. *Botanical Bulletin: Academia Sinica Taipei*, *27*, 247–253.
- Peralta, I., Rodriguez, J. G., & Arroyo, M. T. K. (1992). Breeding system and aspects of pollination in Acacia caven (Mol.) Mol. (Leguminosae: Mimosoideae) in the mediterranean-type climate zone of central Chile. Botanische Jahrbürcher Fur Systematik, Pflanzengeschichte Und Pflanzengeographie, 114(3).
- Percival, M. (1974). Floral ecology of coastal scrub in southeast Jamaica. *Biotropica*, 104–129.
- Philp, J., & Sherry, S. (1946). The degree of natural crossing in green wattle, *Acacia decurrens* Willd. And its bearing on wattle breeding. *Journal of the South African Forestry Association*, 14(1), 1–28.
- Piechowski, D. (2007). *Reproductive ecology, seedling performance, and population structure of* Parkia pendula *in an Atlantic forest fragment in Northeastern Brazil* [PhD Thesis]. Universität Ulm.
- Pieters, A. (1926). Self-sterility or fertility in *Trifolium* and *Melilotus*. *Memoirs, Horticultural Society of New York, 3*, 285–289.

- Pires, I., Andrade, G. de C., & Araujo, M. de S. (1986). Genetic variation for growth characteristics in *P. juliflora* progenies. *International Conference on Prosopis*.
- Plants for a Future. (2014). In *Gymnocladus chinensis Baill*. https://pfaf.org/user/Plant.aspx ?LatinName=Gymnocladus+chinensis
- Platt, W. J., Hill, G. R., & Clark, S. (1974). Seed production in a prairie legume (*Astragalus canadensis* L.). *Oecologia*, *17*(1), 55–63.
- Pojar, J. (1974). Reproductive dynamics of four plant communities of southwestern British Columbia. *Canadian Journal of Botany*, *52*(8), 1819–1834.
- Pometti, C., Cialdella, A., Vilardi, J., & Saidman, B. (2007). Morphometric analysis of varieties of *Acacia caven* (Leguminosae, Mimosoideae): Taxonomic inferences in the context of other Argentinean species. *Plant Systematics and Evolution*, 264(3), 239–249.
- Pometti, C. L., Bessega, C. F., Vilardi, J. C., & Saidman, B. O. (2013). Comparison of mating system parameters and genetic structure in three natural scenarios of *Acacia visco* (Leguminosae, Mimosoideae). *Plant Systematics and Evolution*, 299(4), 761–771.
- Prata de Assis Pires, J., & Freitas, L. (2008). Reproductive biology of two tree species of Leguminosae in a Montane Rain Forest in southeastern Brazil. *Flora Morphology, Distribution, Functional Ecology of Plants*.
- Pritchard, A., & t'Mannetje, L. (1967). The breeding systems and some interspecific relations of a number of African *Trifolium* spp. *Euphytica*, *16*(3), 324–329.
- Purseglove, J. W., & others. (1968). Tropical crops. Dicotyledons 1 and 2. *Tropical Crops. Dicotyledons* 1 and 2.
- Putiyevsky, E. (1970). Chromosome number and genetic systems in several *Trifolium* species related to *T. alexandrinum*. *Chromosoma*, *30*, 476–482.
- Queiroz, L. P. de, da Silma, M., Ramos, A. K. B., & Pizarro, E. A. (1997). Estudos reprodutivos em Cratylia argentea (Desv.) O. Kuntze e Cratylia mollis Mart. Ex Benth. (Leguminosae-Papilionoideae). 19(3), 20–23.
- Rabasa, S. G., Gutiérrez, D., & Alcántara, A. E. (2009). Breeding system of the Iberian endemic shrub *Colutea hispanica* (Leguminosae). *Anales Del Jardín Botánico de Madrid*, 66(2), 279–284.
- Raimondo, F. M., & Domina, G. (2007). Two new *Mimosaceae* naturalized in Italy. *Flora Mediterranea*, *17*, 209–216.
- Rambuda, T., & Johnson, S. (2004). Breeding systems of invasive alien plants in South Africa: Does Baker's rule apply? *Diversity and Distributions*, *10*(5-6), 409–416.
- Ramirez, N., & Brito, Y. (1990). Reproductive biology of a tropical palm swamp community in the Venezuelan Ilanos. *American Journal of Botany*, *77*(10), 1260–1271.
- Ramirez, N., Sobrevila, C., de Enrech, N. X., & Ruiz-Zapata, T. (1984). Floral biology and breeding system of *Bauhinia benthamiana* Taub. (Leguminosae), a bat-pollinated tree in Venezuelan Llanos. *American Journal of Botany*, *71*(2), 273–280.
- Rao, H. (1954). Pollination mechanisms in Acacia catechu. Journal of the Indian Botanical Society, 33, 93–97.
- Rao, S. P., & Raju, A. S. (2002). Pollination ecology of the Red Sanders *Pterocarpus santalinus* (Fabaceae), an endemic and endangered tree species. *Current Science*, *83*(9), 1144–1148.

- Rathcke, B. J., & Kass, L. B. (2003). Floral biology of *Calliandra haematomma var. correllii* (Fabaceae), an andromonoecious bird-pollinated shrub endemic to the Bahamas. *Proceedings of the Ninth Symposium on the Natural History of the Bahamas*, 21–27.
- Real, D., Dalla Rizza, M., Reyno, R., & Quesenberry, K. H. (2007). Breeding system of the aerial flowers in an amphicarpic clover species: *Trifolium polymorphum*. *Crop Science*, *47*(4), 1401–1406.
- Ribeiro, R. A., & Lovato, M. B. (2004). Mating system in a neotropical tree species, *Senna multijuga* (Fabaceae). *Genetics and Molecular Biology*, 27(3), 418–424.
- Richards, K. (1986). Pollination requirements of cicer milkvetch, Astragalus cicer L. Journal of Range Management, 39(5), 457–459.
- Rivera-Hernández, K. L., Torres-González, A. M., Schultze-Kraft, R., & Peters, M. (2011). Biología floral del ecotipo Yapacaní de Cratylia aff. argentea (Fabaceae).
- Roberts, D. J., Werner, D. J., Wadl, P. A., & Trigiano, R. N. (2015). Inheritance and allelism of morphological traits in eastern redbud (*Cercis canadensis* L.). *Horticulture Research*, *2*, 15049.
- Robertson, A. W., Kelly, D., & Ladley, J. J. (2011). Futile selfing in the trees *Fuchsia excorticata* (On-agraceae) and *Sophora microphylla* (Fabaceae): Inbreeding depression over 11 years. *International Journal of Plant Sciences*, 172(2), 191–198.
- Rocha, O. J., & Aguilar, G. (2001). Variation in the breeding behavior of the dry forest tree *Enterolobium* cyclocarpum (Guanacaste) in Costa Rica. American Journal of Botany, 88(9), 1600–1606.
- Rodger, J. G., & Johnson, S. D. (2013). Self-pollination and inbreeding depression in *Acacia dealbata*: Can selfing promote invasion in trees? *South African Journal of Botany*, *88*, 252–259.
- Rodriguez-Riano, T., Ortega-Olivencia, A., & Devesa, J. (1999). Reproductive biology in two Genisteae (Papilionoideae) endemic of the western Mediterranean region: Cytisus striatus and Retama sphaerocarpa. Canadian Journal of Botany, 77(6), 809–820.
- Rodríguez-Riaño, T., Ortega-Olivencia, A., & Devesa, J. A. (2004). Reproductive biology in *Cytisus multiflorus* (Fabaceae). *Annales Botanici Fennici*, 41, 179–188.
- Roy, A., Malaviya, D., & Kaushal, P. (2005). Pollination behaviour among different breeding populations of Egyptian clover. *Plant Breeding*, *124*(2), 171–175.
- Rymer, P., Morris, E., & Richardson, B. (2002). Breeding system and population genetics of the vulnerable plant *Dillwynia tenuifolia* (Fabaceae). *Austral Ecology*, *27*(3), 241–248.
- Sahai, K. (2009). Reproductive biology of two species of Canavalia DC.(Fabaceae)A non-conventional wild legume. *Flora-Morphology, Distribution, Functional Ecology of Plants, 204*(10), 762–768.
- Sajjad, A., Saeed, S., & Bashir, M. A. (2012). Spatial variation in pollinator communities and reproductive performance of *Prosopis juliflora* (Fabaceae). *Journal of Pollination Ecology*, 8(9), 59–66.
- Salinas, A. O. D., & Sánchez, M. S. (1977). Biologia floral del genero *Cassia* en la region de Los Tuxtlas, Veracruz. *Boletín*, *37*, 5.
- Santos, G., Machado, V., & Giannotti, E. (1993). Visitantes florais de Bauhinia variegata L. (Leguminosae, Caesalpiniaceae). Revista Brasileira de Entomolgia, 37, 737-750.
- Saraiva, L., Cesar, O., & Monteiro, R. (1996). Breeding systems of shrubs and trees of a Brazilian savanna. *Brazilian Archives of Biology and Technology*, 39(4), 751–763.
- Sarala, B., Lokesha, R., & Vasudeva, R. (1999). Anther dimorphism, differential anther dehiscence, pollen viability and pollination success in *Caesalpinia pulcherrimma* L. (Fabaceae). *Current Science*, 76(11), 1490–1494.

Sauer, J. (1964). Revision of Canavalia. Brittonia, 106-181.

- Schoen, D. J., & Brown, A. H. (1991). Whole-and part-flower self-pollination in *Glycine clandestina* and *G. argyrea* and the evolution of autogamy. *Evolution*, *45*(7), 1651–1664.
- Seavey, S. R., & Bawa, K. S. (1986). Late-acting self-incompatibility in angiosperms. *The Botanical Review*, *52*(2), 195–219.
- Sedgley, M., Harbard, J., Smith, R., Wickneswari, R., & Griffin, A. (1992). Reproductive biology and interspecific hybridization of *Acacia mangium* and *Acacia auriculiformis* A. Cunn. Ex Benth (Leguminosae, Mimosoideae). *Australian Journal of Botany*, 40(1), 37–48.
- Severns, P. M., & Lewis, B. S. (2007). Is it appropriate to rely on seed set to assess candidate plant populations for genetic rescue? A case study with a threatened species. *Natural Areas Journal*, 27(4), 313–319.
- Shi, X., Wang, J. C., Zhang, D. Y., Gaskin, J., & Pan, B. R. (2010). Pollination ecology of the rare desert species *Eremosparton songoricum* (Fabaceae). *Australian Journal of Botany*, *58*(1), 35–41.
- Sigrist, M. R., Stefanello, T. H., de Souza, C. S., Vargas, W., Almeida, K. S. M., Laroca, S., & Mansano, V.
  F. (2018). Phenology and pollination ecology of *Prosopis rubriflora* (Leguminosae, Mimosoideae), a species from the semi-arid Brazilian Chaco. *Brazilian Journal of Botany*, *41*(1), 103–115.
- Silvertown, J., McConway, K. J., Hughes, Z., Biss, P., Macnair, M., & Lutman, P. (2002). Ecological and genetic correlates of long-term population trends in the park grass experiment. *The American Naturalist*, *160*(4), 409–420.
- Simpson, B. B. (1977). Breeding systems of dominant perennial plants of two disjunct warm desert ecosystems. *Oecologia*, *27*(3), 203–226.
- Sinskaya, E. (1961). Flora of the cultivated plants of the USSR XIII. Perennial Leguminous Plants. Part I: Medicago, Sweet Clover, Fenugreek. Israel Programme for Scientific Translations, Jerusalem, 514–570.
- Skolmen, R. G. (2015). Eastern Region State and Private Forestry. In Prosopis pallida (Humb. & Bonpl. ex Willd.) H.B.K. http://www.na.fs.fed.us/pubs/silvics\_manual/volume\_2/prosopis/ pallida.htm
- Smartt, J. (1964). Cross-compatibility relationship between The cultivated peanut *Arachis hypogaea* L. And other species of the genus *Arachis. Ph. D. Thesis, North Carolina State University*.
- Smitha, G., & Thondaiman, V. (2016). Reproductive biology and breeding system of *Saraca asoca* (Roxb.) De Wilde: A vulnerable medicinal plant. *SpringerPlus*, *5*(1), 2025.
- Soloman Raju, A. J., & Purnachandra Rao, S. (2002). Reproductive ecology of *Acacia concinna* (Shikakai) and Smecarpus anacardium (marking nut), with a note on pollinator conservation in the eastern ghats of Visakhapatnam district, Andhra Pradesh. *National Seminar on Conservation of Eastern Ghats*, 57–58.
- Sorensson, C. T., & Brewbaker, J. L. (1994). Interspecific compatibility among 15 *Leucaena* species (Leguminosae: Mimosoideae) via artificial hybridizations. *American Journal of Botany*, *81*(2), 240–247.
- Sousa, D. M. M., do Rêgo, M. M., Meira, R. M. S. A., & do Rêgo, E. R. (2016). Behavior of the pollen tube of *Poincianella pyramidalis* (Tul.) LP Queiroz after compatible and incompatible crosses. *African Journal of Agricultural Research*, 11(25), 2193–2199.
- Spellenberg, R., Earle, C. J., & Nelson, G. (2014). *Trees of Western North America*. Princeton University Press.

- Stace, H. (1984). Genetic Systems in *Stylosanthes*. In *The Biology and Agronomy of Stylosanthes*. Academic Press Australia.
- Stephenson, A. G., & Winsor, J. A. (1986). Lotus corniculatus regulates offspring quality through selective fruit abortion. Evolution, 40(3), 453–458.
- Stirton, C. H. (1975). A contribution to the knowledge of the genus Eriosema (Leguminosae, Lotoideae) in Southern Africa (excluding Mozambique and Rhodesia) [PhD thesis].
- Sugden, E. A. (1985). Pollinators of *Astragalus monoensis* Barneby (Fabaceae): New host records; potential impact of sheep grazing. *The Great Basin Naturalist*, *45*(2), 299–312.
- Swoboda, K. A. (2007). The pollination ecology of Hedysarum boreale Nutt. (Fabaceae) and evaluation of its pollinating bees for restoration seed production [Master's Thesis]. Utah State University.
- Sørensen, M. (1996). Yam Bean: Pachyrhizus DC.-Promoting the conservation and use of underutilized and neglected crops. 2. (Vol. 2). Bioversity International.
- Tandon, R., Shivanna, K., & Ram, H. (2001). Pollination biology and breeding system of Acacia senegal. Botanical Journal of the Linnean Society, 135(3), 251–262.
- Tandon, R., Shivanna, K., & Ram, H. (2003). Reproductive biology of *Butea monosperma* (Fabaceae). Annals of Botany, 92(5), 715–723.
- Tanner, D. A., Clark, C., & Pitts, J. P. (2013). Pollination biology of Astragalus phoenix (Fabaceae) with notes on the natural history of its pollinator, Anthophora porterae (Hymenoptera: Apidae). Western North American Naturalist, 73(3), 373–381.
- Team, R. E. A. (2012). *Peirson's Milk-Vetch* (Astragalus magdalenae *var.* peirsonii) [Baseline Biology Report]. Dudek Interational.
- Tedesco, S., Dall'Agnol, M., Schifino-Wittmann, M., & Valls, J. (2000). Mode of reproduction of Brazilian species of *Adesmia* (Leguminosae). *Genetics and Molecular Biology*, 23(2), 475–478.
- Teixeira, S. de P., Pereira, R. A. S., & Ranga, N. T. (2006). Components of fecundity and abortion in a tropical tree, *Dahlstedtia pentaphylla* (Leguminosae). *Brazilian Archives of Biology and Technology*, 49(6), 905–913.
- Teixeira, S., Prakash, N., & Ranga, N. (2001). Ovule and early seed development related to seed abortion in *Dahlstedtia pinnata* and *D. pentaphylla* (Leguminosae, Papilionoideae). *Phytomorphology*, 51(1), 41–50.
- Thimmaiah, M. R., Choudhary, S. B., Sharma, H. K., Kumar, A. A., Bhandari, H., Mitra, J., & Karmakar, P. G. (2018). Late-acting self-incompatibility: A barrier to self-fertilization in sunnhemp (*Crotalaria juncea* L.). *Euphytica*, 214(2), 19.
- Thompson, J. (1931). Studies in advancing sterility. Part V. The theory of the leguminous strobilus. *Publications of the Hartley Botanical Laboratories, University of Liverpool, 7*, 1–79.
- Thomson, J. (1938). Cross- and self-fertility in sainfoin. Annals of Applied Biology, 25(4), 695-704.
- Todd, F. E., & Vansell, G. (1952). The role of polllinating insects in legume-seed production. *Proc. 6th Int. Grassld Congr.*
- Townsend, C. (1970). Phenotypic diversity for agronomic characters and frequency of self-compatible plants in *Trifolium ambiguum*. *Canadian Journal of Plant Science*, *50*(3), 331–338.
- Trapp, E. J., & Hendrix, S. D. (1988). Consequences of a mixed reproductive system in the hog peanut, *Amphicarpaea bracteata*, (Fabaceae). *Oecologia*, *75*(2), 285–290.

- Tropical Forages. (n.d.). In *Stylosanthes capitata*. https://www.tropicalforages.info/text/ entities/stylosanthes\_capitata.htm
- Tsai, L. M. (1988). Studies on *Acacia mangium* in Kemasul Forest, Malaysia. I. Biomass and productivity. *Journal of Tropical Ecology*, 4(03), 293–302.
- Tybirk, K. (1993). Pollination, breeding system and seed abortion in some African Acacia. Botanical Journal of the Linnean Society, 112(2), 107–137.
- Ufer, M. R. (1930). Die verwendung von bienen bei kreuzungsversuchen mit steinklee (*Melilotus*). Der Züchter, 2, 305–308.
- Uphof, J. T. (1938). Cleistogamic flowers. The Botanical Review, 4(1), 21-49.
- Valtueña, F. J., Ortega-Olivencia, A., Rodriguez-Riaño, T., & Lopez, J. (2008). Reproductive biology in Anagyris foetida L. (Leguminosae), an autumnWinter flowering and ornithophilous Mediterranean shrub. Botanical Journal of the Linnean Society, 157(3), 519–532.
- Valtueña, F. J., Rodríguez-Riaño, T., Espinosa, F., & Ortega-Olivencia, A. (2010). Self-sterility in two *Cytisus* species (Leguminosae, Papilionoideae) due to early-acting inbreeding depression. *American Journal of Botany*, 97(1), 123–135.
- Van Wyk, B. (1990). Studies in the genus *Lotononis* (Crotalarieae, Fabaceae). 13. Two new species and notes on the occurrence of cleistogamy in the section Leptis. *Bothalia*, *20*(1), 17–22.
- Venkatesh, C. S. (1994). Conservation of forest genetic resources. In S. Deshmukh & V. Balaji (Eds.), Conservation of Mangrove Forest Genetic Resources : A Training Manual (pp. 77–83). M.S. Swaminathan Research Foundation.
- Venter, H. (1996). Difficulties with cross-pollinating five cowpea lines: Technique development. In *The Biodiversity of African Plants* (pp. 656–660). Kluwer Academic Publishers.
- Venturieri, G. (1996). Reproductive ecology of Schizolobium amazonicum Huber ex Ducke and Sclerolobium paniculatum Vogel (Caesalpiniaceae) and its importance in forestry management projects. ISHS Acta Horticulturae 437, 65–70.
- Villalobos, S., & Ramírez, N. (2010). Biologia reproductiva de *Crotalaria micans* Link (Fabaceae): Especie colonizadora de amplia distribucion geographica. *Acta Botánica Venezuelica*, *33*, 67–81.
- Vivarelli, D., Petanidou, T., Nielsen, A., & Cristofolini, G. (2011). Small-size bees reduce male fitness of the flowers of *Ononis masquillierii* (Fabaceae), a rare endemic plant in the northern Apennines. *Botanical Journal of the Linnean Society*, 165(3), 267–277.
- Wagner, D. (2000). Pollen viability reduction as a potential cost of ant association for *Acacia constricta* (Fabaceae). *American Journal of Botany*, *87*(5), 711–715.
- Wang, Y. J., Nair, R. M., Mu, C. S., & Dundas, I. S. (2011). Floral morphology and pollination system in the native Australian perennial pasture legume *Cullen australasicum* (syn. *Psoralea australasica*). *Crop and Pasture Science*, 61(12), 1001–1008.
- Watrous, K. M. (2010). *Reproductive ecology of Astragalus filipes, a great basin restoration legume* [Master's Thesis]. Utah State University.
- Watrous, K. M., & Cane, J. H. (2011). Breeding biology of the threadstalk milkvetch, *Astragalus filipes* (Fabaceae), with a review of the genus. *The American Midland Naturalist*, *165*(2), 225–240.
- Whitney, K. D. (2005). Evidence for simple genetic control of a fruit-colour polymorphism in *Acacia ligulata*. *Australian Journal of Botany*, *53*(4), 363–366.
- Williams, I. H. (1987). The pollination of lupins. *Bee World*, 68(1), 10–16.

- Williams, W. (1951). Genetics of incompatibility in alsike clover, *Trifolium hybridum*. *Heredity*, *5*(1), 51–73.
- Windler, D. (1966). A revision of the genus *Neptunia* (Leguminosae). *Australian Journal of Botany*, 14(3), 379–420.
- Windler, D. (1973). Field and garden studies in *Crotalaria sagittalis* L. And related species. *Phytologia*. Wojciechowska, B. (1972). Pollination and fertilization in *Ornithopus* sp. *Genet Polon*.
- Wolowski, M., & Freitas, L. (2010). Sistema reprodutivo e polinização de *Senna multijuga* (Fabaceae) em mata atlântica montana. *Rodriguésia*, *61*(2), 167–179.
- Wood, H. (2010). Tropical Trees for Life. http://www.treesforlife.info/gmptsf/acacia-senegal .htm
- Young, A., & Brown, A. (1998). Comparative analysis of the mating system of the rare woodland shrub *Daviesia suaveolens* and its common congener *D. mimosoides*. *Heredity*, *80*, 374–381.
- Zandstra, I. I., & Grant, W. F. (1968). The biosystematics of the genus *Lotus* (Leguminosae) in Canada. I. Cytotaxonomy. *Canadian Journal of Botany*, *46*(5), 557–583.
- Zapata, T. R., & Arroyo, M. T. K. (1978). Plant reproductive ecology of a secondary deciduous tropical forest in Venezuela. *Biotropica*, *10*(3), 221–230.
- Zenni, R., Wilson, J., Le Roux, J., & Richardson, D. (2009). Evaluating the invasiveness of *Acacia para*doxa in South Africa. South African Journal of Botany, 75(3), 485–496.
- Zohary, D. (2002). Domestication of the carob (*Ceratonia siliqua* L.). *Israel Journal of Plant Sciences*, 50(sup1), 141–145.

## A.0.5 Supplemental figures and tables

TABLE XI: **Summary of studies examining the genetic phenomena possibly related to selfincompatibility (SI) in Fabaceae.** We found nine studies reporting work on the genetic inheritance pattern of differential fruit set following self- and cross-fertilization in the family.

Study	Species	Findings	
Miri & Bubar, 1965	Lotus corniculatus	Report disomic inheritance of S-alleles at two location	
		in incompatible progeny and tetrasomic inheritance	
		of S-alleles at one loci in selfed progeny.	
Lundqvist, 1993	Lotus tenuis	Performed reciprocal crosses and reported multi-	
		locus control with asymmetric directional cross fail-	
		ures, with "complex genetic control."	
Gettys & Johnson, 1944	Melilotus officinalis	Results from reciprocal crossing experiments indicate	
		that SC is dominant to "self-sterility."	
Sandal, 1951	Melilotus officinalis	Wide segregation for self-compatibility occurred in S1	
		and S2 progenies and in F1 backcross progenies, sug-	
		gesting SC is conditioned by more than one locus.	
Williams, 1951	Trifolium hybridum	Crosses indicated an S-locus with "extensive" sterility	
		alleles.	
Brewbaker, 1955	Trifolium negrescens	Diallelic crosses suggest that SI response is under	
		control of oppositional S-alleles of the gametophytic	
		type.	
Rinke & Johnson, 1941	Trifolium pratense	Report the existence of a single sterility allele.	
Atwood, 1942	Trifolium repens	Crosses indicate a series of oppositional allomorphs	
		(S-alleles), consistent with gametophytic SI.	
Casey et al., 2010	Trifolium repens	Found a single S-locus with four different S-	
		haplotypes segregating in a population, with	
		haplotypes from both parents being transmitted in	
		equal proportions.	

Species	Site of Inhibition	Study
Acacia retinodes	Ovule	Kenrick & Knox 1989
Acacia senegal	Post-fertilization (entered embryo sac but no fruits formed)	Tandon & Shiyanna 2001
Bauhinia forficata	Style (upper)	Neto 2013
Caesalninia calveina	Ovary	Lewis & Gibbs et al. 1999 (in Gibbs 2014)
Caesalninia nyramidalis	Ovary	Borges et al. 2008 (in Gibbs 2014)
Calliandra calothyrsus	Ovary	Chamberlain 1998 (in Gibbs 2014)
Crotalaria juncea	Style (slower growth) & ovary	Thimmaiah et al. 2018
Cytisus multiflorus	Post-fertilization (seed abortion)	Valtuena et al. 2010
Cytisus sconarius	Style stigma (weak)	Parker & Haubensak 2002
Cytisus striatus	Post-fertilization (seed abortion)	Valtuena et al 2010
Dalhergia retusa	Ovule	Seavey & Bawa 1986
Dintervx nanamensis	Ovule	Seavey & Bawa 1986
Frythring spn	Style (slowed growth) post-fertilization (fruit abortion)	Neill 1988
Genista monospessulana	Style stigma (weak)	Parker & Haubensak 2002
Hedysarum coronarium	Style	Prados 1988 (in Heras et al. 2001)
Hedysarum spinosissimum (Hedysarum alomeratum)	Stigma (full compatibility with stigmatic membrane disruption)	Heras et al. 2001
Hedysarum humile	Style	Prados 1988 (in Heras et al. 2001)
Hymenaea stigonocarna	Post-fertilization (floral abscission, reduced ovule size)	Gibbs et al. 1999
Inga sierrae (Inga hrenesii)	Ovule	Koptur 1984
Leucaena leucocenhala	Stigma selective embryo abortion	Ganeshajah et al. 1986
Lonchocarnus sn		Bawa et al. 1985
Lonchocarpus sp.	Ovule	Dobrofsky & Grant 1980
Lotus corniculatus	Stigma (slight compatibility with stigmatic membrane disruption)	Bubar 1958
Lotus conneniatus	Stigma (single compatibility with stigmatic membrane disruption)	Bubar 1958
Lotus saureotens	Stigma (slight compatibility with stigmatic membrane disruption)	Bubar 1958
Lotus pedunculatus (Lotus uliginosus)	Stigma (slight compatibility with stigmatic disruption)	Bubar 1958
Lotus weilleri	Stigma (signe compatibility with stigmatic membrane disruption)	Bubar 1958
Medicago sativa	Style (slowed growth) ovule	Brink 1938 & 1939
Melilotus officinalis	Stigma style (slowed growth)	Brink 1934
Phaseolus coccineus	Style ovule post-fertilization	Hawkins & Evans 1973
Phaseolus vulgaris	Style ovule post-fertilization	Hawkins & Evans, 1973
Cenostiama pyramidale	Style	Sousa et al. 2016
Prosonis flexuosa	Ovule post-fertilization	Masuellii 1989
Pseudonintadenia contorta	Ovary	Prata de Assis Pires & Frietas 2008
Pterocarnus macrocarnus	Post-fertilization (ovule and fruit abortion)	Doungvotha & Owens 2002
Senna multijuga	Post-fertilization (complete uniform pistil abortion)	Wolowski & Freitas 2010
Senna svivestris	Post-fertilization (complete)	Carvalho & Oliveira 2003
Strongylodon macrobotrys	Ovary	Prychid et al. 1998 (in Gibbs 2014)
Swartzia apetala	Ovary	Chiara Moco. 1996 (in Gibbs, 2014)
Tamarindus indica	Style (slowed growth) ovary	Diallo et al. 2008
Trifolium granduliferum	Stigma (no pollen tube growth)	Nair et al., 2007
Trifolium pratense	Style (slowed growth)	Leduc. 1990: Williams, 1933: Pieters, 1927
Trifolium pratense	Stigma (no pollen tube growth without stigmatic membrane disruption)	Heslop-Harrison, 1983
Trifolium repens	Stigma, style	Atwood, 1940
Trifolium repens	Style (mid)	Leduc. 1990
Tylosema esculentum	Style (heterostylous)	Hartley et al., 2002
		1

TABLE XII: **Summary of studies examining site-of-fertilization failure following self-pollination in Fabaceae species.** We conservatively include apparently post-zygotic phenomena, which may not constitute SI responses, but comprise the diverse situations sometimes described as late-acting SI. Currently-accepted species binomial names are reported. If names differ from the original study, the original names are listed in parentheses. Findings are summarized in the SITE OF INHIBITION column.



Figure 22: **Proportion of legume species by genus that are SI.** Numbers in parenthesis next to each genus indicate the number of species used to calculate the proportion. Data is available for 163 genera.



Figure 23: **Species-level Fabaceae phylogeny displaying SI and SC species (number of species = 502).** This cladogram is modified from a Fabales tree found in Smith & Brown (2018). We pruned their tree to retain only species with breeding system data; see main article text for details.

#### Appendix **B**

#### **APPENDIX B**

#### **B.0.1** Collected works for Orchidaceae breeding system database

- Abdala-Roberts, L., Parra-Tabla, V., & Navarro, J. (2007). Is floral longevity influenced by reproductive costs and pollination success in *Cohniella ascendens* (Orchidaceae)? *Annals of Botany*, 100(6), 1367– 1371.
- Ackerman, J. D. (1975). Reproductive biology of Goodyera oblongifolia (Orchidaceae). Madroño, 23(4), 191–198.
- Ackerman, J. D. (1976). *Biosystematic studies in the genus* Piperia (*Orchidaceae*) [PhD thesis]. Humboldt State University.
- Ackerman, J. D. (1989). Limitations to sexual reproduction in *Encyclia krugii* (Orchidaceae). Systematic Botany, 101–109.
- Ackerman, J. D. (1995). Orchid flora of Puerto Rico and the Virgin Islands. New York Botanical Garden.
- Ackerman, J. D., & Mesler, M. R. (1979). Pollination biology of *Listera cordata* (Orchidaceae). *American Journal of Botany*, 66(7), 820–824.
- Ackerman, J. D., & Montalvo, A. M. (1990). Short-and long-term limitations to fruit production in a tropical orchid. *Ecology*, *71*(1), 263–272.
- Ackerman, J. D., Morales, M., & Tremblay, R. (2011). Darwin's orchids: Their variation, plasticity, and natural selection. *Lankesteriana International Journal on Orchidology*, *11*(3), 179–184.
- Ackerman, J. D., & Oliver, J. M. (1985). Reproductive biology of *Oncidium variegatum*: Moon phases, pollination, and fruit set. *American Orchid Society Bulletin*, 54(3), 326–329.
- Adams, R., & Goss, G. (1976). Reproductive biology of the epiphytic orchids of Florida. III. *Epidendrum* anceps Jacquin. American Orchid Society Bulletin.
- Agnew, J. (1986). Self-compatibility/incompatibility in some orchids of the subfamily Vandoideae. *Plant Breeding*, *97*(2), 183–186.
- Aguiar, J. M., Pansarin, L. M., Ackerman, J. D., & Pansarin, E. R. (2012). Biotic versus abiotic pollination in *Oeceoclades maculata* (Lindl.) Lindl. (Orchidaceae). *Plant Species Biology*, *27*(1), 86–95.
- Aguiar, J. M. R. B. V., & Pansarin, E. R. (2019). Deceptive pollination of *Ionopsis utricularioides* (Oncidiinae: Orchidaceae). *Flora*, 250, 72–78.
- Albores-Ortiz, O., & Sosa, V. (2006). Polinización de dos especies simpátricas de *Stelis* (Pleurothallidinae, Orchidaceae). *Acta Botánica Mexicana, 74*, 155–168.
- Alexandersson, R., & Ågren, J. (1996). Population size, pollinator visitation and fruit production in the deceptive orchid *Calypso bulbosa*. *Oecologia*, *107*(4), 533–540.
- Ames, O., & Correll, D. S. (1953). Orchids of Guatemala.

- Amorim, F. W., Wyatt, G. E., & Sazima, M. (2014). Low abundance of long-tongued pollinators leads to pollen limitation in four specialized hawkmoth-pollinated plants in the Atlantic Rain Forest, Brazil. *Naturwissenschaften*, 101(11), 893–905.
- Aragon, S., & Ackerman, J. D. (2001). Density effects on the reproductive success and herbivory of *Malaxis massonii. Lindleyana*, *16*(1), 3–12.
- Argue, C. L. (2011a). The pollination biology of North American orchids: Volume 1: North of Florida and Mexico (Vol. 1). Springer Science & Business Media.
- Argue, C. L. (2011b). The pollination biology of North American orchids: Volume 2: North of Florida and Mexico (Vol. 2). Springer Science & Business Media.
- Barbosa, A. A. A. (1997). *Biologia reprodutiva de uma comunidade de campo sujo, Uberlândia/MG* [PhD Thesis]. Universidade Estadual de Campinas.
- Barbosa, A. R., de Melo, M. C., & Borba, E. L. (2009). Self-incompatibility and myophily in *Octomeria* (Orchidaceae, Pleurothallidinae) species. *Plant Systematics and Evolution*, *283*(1), 1–8.
- Barrett, S. C., & Helenurm, K. (1987). The reproductive biology of boreal forest herbs. I. Breeding systems and pollination. *Canadian Journal of Botany*, 65(10), 2036–2046.
- Bartareau, T. (1994). The reproductive ecology of *Dendrobium toressae* (Bailey) Dockr., A geographically restricted species in north-east Queensland. *The Orchadian (Australasian Native Orchid Society)*, 11, 106–111.
- Bateman, R. M., Sramkó, G., & Rudall, P. J. (2015). Floral miniaturisation and autogamy in boreal-arctic plants are epitomised by Iceland's most frequent orchid, *Platanthera hyperborea*. *PeerJ*, *3*, e894.
- Bates, R. (1978). Pollination of orchids: Part 9. *Journal of the Native Orchid Society of South Australia,* 2, 7–8.
- Bates, R. (1984). Australia's colourful sun-orchids: Thelymitra. Australian Orchid Review, 49, 109-101.
- Beardsell, D., Clements, M., Hutchinson, J., & Williams, E. (1986). Pollination of *Diuris maculata* R Br (Orchidaceae) by floral mimicry of the native legumes *Daviesia spp* and *Pultenaea scabra* R Br. *Australian Journal of Botany*, 34(2), 165–173.
- Bender, J. (1985). The reproductive-biology of *Cypripedium candidum* Muhl ex Willd (Orchidaceae). *Ohio Journal of Science*, *85*, 12–12.
- Benson, D., & McDougall, L. (2005). Ecology of sydney plant species. Part 10. Monocotyledon families Lemnaceae to Zosteraceae. *Cunninghamia*, 9(1), 16–212.
- Bernhardt, P., & Burns-Balogh, P. (1986). Floral mimesis in *Thelymitra nuda* (Orchidaceae). *Plant Systematics and Evolution*, 151(3), 187–202.
- Berry, P. E., & Calvo, R. N. (1991). Pollinator limitation and position dependent fruit set in the high Andean orchid *Myrosmodes cochleare* (Orchidaceae). *Plant Systematics and Evolution*, 174(1), 93– 101.
- Bishop, A. (1996). *Field guide to the orchids of New South Wales and Victoria*. University of New South Wales Press Sydney, Australia.
- Blood, K. (2001). Disa bracteata. In Environmental weeds: A field guide for SE Australia. CH Jerram & Associates.
- Borba, E. L., Barbosa, A. R., de Melo, M. C., Gontijo, S. L., & de Oliveira, H. O. (2011). Mating systems in the Pleurothallidinae (Orchidaceae): Evolutionary and systematic implications. *Lankesteriana International Journal on Orchidology*, 11(3), 207–221.

- Borba, E. L., Semir, J., & Shepherd, G. J. (2001). Self-incompatibility, inbreeding depression and crossing potential in five Brazilian *Pleurothallis* (Orchidaceae) species. *Annals of Botany*, 88(1), 89–99.
- Borba, E. L., Shepherd, G. J., & Semir, J. (1999). Reproductive systems and crossing potential in three species of *Bulbophyllum* (Orchidaceae) occurring in Brazilian "campo rupestre" vegetation. *Plant Systematics and Evolution*, 217(3), 205–214.
- Bory, S., Lubinsky, P., Risterucci, A.-M., Noyer, J.-L., Grisoni, M., Duval, M.-F., & Besse, P. (2008). Patterns of introduction and diversification of *Vanilla planifolia* (Orchidaceae) in Reunion Island (Indian Ocean). *American Journal of Botany*, 95(7), 805–815.
- Bouetard, A., Lefeuvre, P., Gigant, R., Bory, S., Pignal, M., Besse, P., & Grisoni, M. (2010). Evidence of transoceanic dispersion of the genus *Vanilla* based on plastid DNA phylogenetic analysis. *Molecular Phylogenetics and Evolution*, 55(2), 621–630.
- Brian. (2007). Oeceoclades roseovariegata with keikei #4 [Comment]. In Flickr. https://www.flickr .com/photos/aeranthes/2095755021/in/photostream/
- Buragohain, B., Chaturvedi, S., & Puro, N. (2016). Pollination biology of *Phaius tankervilleae* (banks ex L'Herit) Bl.( Orchidaceae). *International Journal of Plant Reproductive Biology*, 8(1), 75–81.
- Burns-Balogh, P. (1984). Evolution of the monandrous Orchidaceae. II. Evolution of characters and pollination in Spiranthoideae. *The Canadian Orchid Journal*, *2*(2), 27–31.
- Burns-Balogh, P., & Bernhardt, P. (1985). Evolutionary trends in the androecium of the Orchidaceae. *Plant Systematics and Evolution*, 149(1), 119–134.
- Bush, S. P., & Beach, J. H. (1995). Breeding systems of epiphytes in a tropical montane wet forest. *Selbyana*, 155–158.
- Cady, L. (1972). Notes on the pollination of Calochilus campestris R. Br. Orchadian, 4, 52-53.
- Calder, D., Adams, P., Slater, A., Williams, E., Knox, R., Gilbert, J., & Bernhardt, P. (1982). The floral biology and breeding system of *Dendrobium speciosum* Sm. *Pollination*, *82*, 184–192.
- Calvo, R. N. (1990). Four-year growth and reproduction of *Cyclopogon cranichoides* (Orchidaceae) in South Florida. *American Journal of Botany*, 77(6), 736–741.
- CaraDonna, P. J., & Ackerman, J. D. (2010). Reproductive assurance for a rewardless epiphytic orchid in Puerto Rico: *Pleurothallis ruscifolia* (Orchidaceae, Pleurothallidinae). *Caribbean Journal of Science*, 46(2), 249–257.
- Catling, P. (1990). Auto-pollination in the Orchidaceae. Orchid Biology. Review and Perspectives. V., 123-158.
- Catling, P. M. (1980). Rain-assisted autogamy in *Liparis loeselii* (L.) LC Rich. (Orchidaceae). *Bulletin of the Torrey Botanical Club*, 525–529.
- Catling, P. M. (1982). Breeding systems of northeastern North American Spiranthes (Orchidaceae). *Canadian Journal of Botany*, *60*(12), 3017–3039.
- Catling, P. M. (1983). Autogamy in eastern Canadian Orchidaceae: A review of current knowledge and some new observations. *Le Naturaliste Canadien*.
- Catling, P. M. (1987). Notes on the breeding systems of *Sacoila lanceolata* (Aublet) Garay (Orchidaceae). *Annals of the Missouri Botanical Garden*, 58–68.
- Charanasri, U. (1977). Self-incompatibility in the Oncidium alliance. Hawaii Orchid Journal, 6, 12–15.
- Charnasri, U. (1973). Chromosome numbers in the genus *Oncidium* and some allied genera. *American Orchid Society Bulletin*, *42*, 518–524.

- Chaturvedi, S. (2009). Mechanism of self pollination in *Cymbidium sinense* (Jacks. Ex Snder). Willd. (Orchidaceae). *The Journal of Plant Reproductive Biology*, 1(1), 1–4.
- Chaturvedi, S. (2012). Unique mode of mechanical autogamy in *Phalaenopsis taenialis* (Lindl.) EA Christensen & Pradhan (Orchidaceae). *The International Journal of Plant Reproductive Biology*, *4*, 1–4.
- Chen, L.-J., Liu, K.-W., Xiao, X.-J., Tsai, W.-C., Hsiao, Y.-Y., Huang, J., & Liu, Z.-J. (2012). The anther steps onto the stigma for self-fertilization in a slipper orchid. *PLoS One*, *7*(5), e37478.
- Cheng, J., Shi, J., Shangguan, F.-Z., Dafni, A., Deng, Z.-H., & Luo, Y.-B. (2009). The pollination of a self-incompatible, food-mimic orchid, *Coelogyne fimbriata* (Orchidaceae), by female *Vespula* wasps. *Annals of Botany*, 104(3), 565–571.
- Christensen, D. E. (1992). Notes on the reproductive biology of *Stelis argentata* Lindl. (Orchidaceae: Pleurothallidinae) in eastern Ecuador. *Lindleyana*, *7*, 28–33.
- Chung, M., & Chung, M. (2005). Pollination biology and breeding systems in the terrestrial orchid Bletilla striata. Plant Systematics and Evolution, 252(1), 1–9.
- Chung, M. Y., & Chung, M. G. (1999). Allozyme diversity and population structure in Korean populations of *Cymbidium goeringii* (Orchidaceae). *Journal of Plant Research*, *112*(2), 139–144.
- Chung, M. Y., & Chung, M. G. (2008). Conservation genetics of the endangered terrestrial orchid *Pogonia minor* in South Korea. *Annales Botanici Fennici*, *45*, 455–464.
- Chung, M. Y., & Chung, M. G. (2010). Population genetic diversity and structure in *Goodyera rosulacea* (Orchidaceae), endemic in Korea, and implications for conservation. *Annales Botanici Fennici*, 47, 460–470.
- Chung, M. Y., & Chung, M. G. (2003). The breeding systems of *Cremastra appendiculata* and *Cymbid-ium goeringii*: High levels of annual fruit failure in two self-compatible orchids. *Annales Botanici Fennici*, 81–85.
- Chung, M. Y., & Park, C.-W. (2008). Fixation of alleles and depleted levels of genetic variation within populations of the endangered lithophytic orchid *Amitostigma gracile* (Orchidaceae) in South Korea: Implications for conservation. *Plant Systematics and Evolution*, *272*(1), 119–130.
- Chung, M. Y., Park, C.-W., & Chung, M. G. (2007). Extremely low levels of allozyme variation in southern Korean populations of the two rare and endangered lithophytic or epiphytic *Bulbophyllum drymoglossum* and *Sarcanthus scolopendrifolius* (Orchidaceae): Implications for conservation. *Biodiversity and Conservation*, 16(3), 775–786.
- Chung, M. Y., Park, C.-W., Myers, E. R., & Chung, M. G. (2007). Contrasting levels of genetic diversity between the common, self-compatible *Liparis kumokiri* and rare, self-incompatible *Liparis makinoana* (Orchidaceae) in South Korea. *Botanical Journal of the Linnean Society*, *153*(1), 41–48.
- Clifford, S., & Owens, S. (1988). Post-pollination phenomena and embryo development in the Oncidiinae (Orchidaceae). In *Sexual Reproduction in Higher Plants* (pp. 407–412). Springer.
- Cole, F. R., & Firmage, D. H. (1984). The floral ecology of *Platanthera blephariglottis*. American Journal of Botany, 71(5), 700–710.
- Comber, J. (1983). The section Calcarifera of the genus *Dendrobium* in Java [Orchids]. *Orchid Digest* (USA).
- Cortez, E. A., Stort, M. N. S., & others. (2010). Pollination of orchids with a pollen mixture. *Naturalia* (*São Paulo*), 33, 127–135.

- Cruger, H. (1865). On the fecundation of Orchids and their morphology. *Botannical Journal of the Linnean Society*, 8.
- Dafni, A. (1987). Pollination in Orchis and related genera: Evolution from reward to deception. Orchid Biology Reviews and Perspectives, 79–104.
- Dafni, A., & Bernhardt, P. (1989). Pollination of terrestrial orchids of Southern Australia and the Mediterranean region: Systematic, ecological, and evolutionary implications. *Evolutionary Biology*, 24, 193– 252.
- Darwin, C. (1862). On the various contrivances by which orchids are fertilised by insects, and on the good effects of intercrossing. John Murray.
- Darwin, C. (1867). Letter 5389A: Darwin, C. R., To Müller, F., 2 February 1867. In F. Burkhardt (Ed.), *The Correspondence of Charles Darwin* (Two thousand, third, Vol. 15). Cambridge University Press.
- Darwin, C. (1869). XVI.Notes On the fertilization of orchids: To the editors of Annals and Magazine of Natural History. *Journal of Natural History*, 4(21), 141–159.
- Darwin, C. (1877). The effects of cross and self fertilisation in the vegetable kingdom. D. Appleton.
- Darwin, C. (1904). Fertilisation of orchids. Murray.
- Datunsolang, T., & Datunsolang, U. (2009). Biologi bunga Vanda dauglas Lindl. (Orchidaceae). Novae Guinea, 1(1), 17-27.
- Davis, R. W. (1986). The pollination biology of Cypripedium acaule (Orchidaceae). Rhodora, 445-450.
- Day, C. (1972). Observations of self-pollination in *Paphiopedilum* species. *American Orchid Society Bulletin.*
- de Melo, M. C., Taucce, P. P. G., & Borba, E. L. (2011). Reproductive biology and isolation mechanisms in rupicolous species of the Acianthera prolifera complex (Orchidaceae) occurring in southeastern Brazil. Plant Systematics and Evolution, 293(1), 161–176.
- de Paiva Neto, V. B., Correa, A. P. M., de Barros, F., Padilha, D. R. C., & Borges, M. C. R. Z. (2015). The *Bletia catenulata* ornamental orchid is self-compatible but pollinator-dependent for reproduction. *Pesquisa Agropecuária Tropical*, *45*(4), 473–479.
- Devy, L., & Sastra, D. R. (2006). Pengaruh radiasi sinar gamma terhadap kultur in vitro tanaman jahe. *Jurnal Sains Dan Teknologi Indonesia*, 8(1).
- Dieringer, G. (1982). The pollination ecology of Orchis spectabilis L. (Orchidaceae). The Ohio Journal of Science, 85(5), 218.
- Dodson, C. H. (1961). Natural pollination of orchids. *Missouri Botanical Garden Bulletin*, 49, 133–152.
- Dressler, R. L. (1968). Pollination by euglossine bees. *Evolution*, 202–210.
- Dressler, R. L. (1981). The orchids: Natural history and classification. Harvard University Press.
- Dressler, R. L. (1970). How to study orchid pollination without any orchids. *American Journal of Botany*, *57*, 749.
- Duque-Buitrago, C. A., Alzate-Quintero, N. F., & Otero, J. T. (2013). Nocturnal pollination by fungus gnats of the Colombian endemic species, *Pleurothallis marthae* (Orchidaceae: Pleurothallidinae). *Lankesteriana International Journal on Orchidology*, *13*(3), 407–417.
- Dutra, D. (2008). *Reproductive biology and asymbiotic seed germination of* Cyrtopodium punctatum, *and endangered orchid* [Master's Thesis]. University of Florida.
- Edens-Meier, R., Westhus, E., Bernhardt, P., & others. (2013). Floral biology of large-flowered *Thelymitra* species (Orchidaceae) and their hybrids in Western Australia. *Telopea*, *15*, 165–183.

- Ehlers, B., Olesen, J., & Ågren, J. (2002). Floral morphology and reproductive success in the orchid *Epipactis helleborine*: Regional and local across-habitat variation. *Plant Systematics and Evolution*, 236(1), 19–32.
- Ellis, A., & Johnson, S. (1999). Do pollinators determine hybridization patterns in sympatric *Satyrium* (Orchidaceae) species? *Plant Systematics and Evolution*, *219*(3), 137–150.
- Emeterio-Lara, A., García-Franco, J. G., Hernández-Apolinar, M., Mora-Herrera, M. E., Toledo-Hernández, V. H., Valencia-Díaz, S., & Flores-Palacios, A. (2018). Endogamy costs and reproductive biology of *Laelia autumnalis*, an endemic orchid of Mexico. *Plant Ecology*, *219*(12), 1423–1434.
- Engler, A. (1912). Das pflanzenreich (Vol. 50). W. Engelmann.
- Fan, X.-L., Barrett, S. C., Lin, H., Chen, L.-L., Zhou, X., & Gao, J.-Y. (2012). Rain pollination provides reproductive assurance in a deceptive orchid. *Annals of Botany*, *110*(5), 953–958.
- Ferguson, C. S., Donham, K., & Brown, J. L. (2005). *Cypripedium fasciculatum* (Orchidaceae) Anthesis and fruit set in relationship to Diapriid activity. *Selbyana*, 103–113.
- Fernald, M. L. (1897). Notes on Florida plants. *Botanical Gazette*, 24(6), 433–436.
- Firmage, D. H., & Cole, F. R. (1988). Reproductive success and inflorescence size of *Calopogon tuberosus* (Orchidaceae). *American Journal of Botany*, *75*(9), 1371–1377.
- Fitzgerald, R. (1888). Australian Orchids (Vol. 1). Charles Porter, Government Printer.
- Forbes, H. O. (1885). On the contrivances for ensuring self-fertilization in some tropical orchids. *Botanical Journal of the Linnean Society*, *21*(138), 538–550.
- Fowlie, J. (1963). Ecological notes: Schomburgkia lyonsii. American Orchid Society Bulletin, 32, 31–33.
- Frei, J. K., OP, & Dodson, C. H. (1972). The chemical effect of certain bark substrates on the germination and early growth of epiphytic orchids. *Bulletin of the Torrey Botanical Club*, 301–307.
- Freyhold, E. von. (1877). Ophrys aranifera. Verhandlungen Des Botanischen Vereins Für Die Provinz Brandenburg, 19, 55–56.
- Fritz, A.-L. (1990). Deceit pollination of *Orchis spitzelii* (Orchidaceae) on the island of Gotland in the Baltic: A suboptimal system. *Nordic Journal of Botany*, *9*(6), 577–587.
- Fritz, A.-L., & Nilsson, L. A. (1996). Reproductive success and gender variation in deceit-pollinated orchids. In *Floral biology* (pp. 319–338). Springer.
- Gale, S. (2007). Autogamous seed set in a critically endangered orchid in Japan: Pollination studies for the conservation of *Nervilia nipponica*. *Plant Systematics and Evolution*, *268*(1), 59–73.
- Galetto, L., Bernardello, G., & Rivera, G. L. (1997). Nectar, nectaries, flower visitors, and breeding system in five terrestrial Orchidaceae from central Argentina. *Journal of Plant Research*, *110*(4), 393–403.
- Gamisch, A., Fischer, G. A., & Comes, H. P. (2014). Recurrent polymorphic mating type variation in Madagascan *Bulbophyllum* species (Orchidaceae) exemplifies a high incidence of auto-pollination in tropical orchids. *Botanical Journal of the Linnean Society*, 175(2), 242–258.
- Gassner, A. (1982). Observations on the pollination of *Diza stolzii* in Southern Malawi. *Orchid Review*, *90*, 230.
- Gigant, R. L., De Bruyn, A., Church, B., Humeau, L., Gauvin-Bialecki, A., Pailler, T., Grisoni, M., & Besse,
  P. (2014). Active sexual reproduction but no sign of genetic diversity in range-edge populations of
  Vanilla roscheri Rchb. F.(Orchidaceae) in South Africa. Conservation Genetics, 15(6), 1403–1415.

- Gijbels, P., Ceulemans, T., Van den Ende, W., & Honnay, O. (2015). Experimental fertilization increases amino acid content in floral nectar, fruit set and degree of selfing in the orchid *Gymnadenia conopsea*. *Oecologia*, 179(3), 785–795.
- Gontijo, S. L., Barbosa, A. R., de Melo, M. C., & Borba, E. L. (2010). Occurrence of different sites of self-incompatibility reaction in four *Anathallis* (Orchidaceae, Pleurothallidinae) species. *Plant Species Biology*, 25(2), 129–135.
- Goss, G. (1977). The reproductive biology of epiphytic orchids of Florida V *Epidendrum difforme* Jacq. *American Orchid Society Bulletin, 46,* 630–636.
- Gregg, K. B. (1975). The effect of light intensity on sex expression in species of *Cycnoches* and *Catasetum* (Orchidaceae). *Selbyana*, 1(2), 101–113.
- Gregg, K. B. (1989). Reproductive biology of the orchid *Cleistes divaricata* (L.) Ames *Var. Bifaria* Fernald growing in a West Virginia meadow. *Castanea*, 57–78.
- Gregg, K. B. (1991). Reproductive strategy of *Cleistes divaricata* (Orchidaceae). *American Journal of Botany*, *78*(3), 350-360.
- Gumbert, A., & Kunze, J. (2001). Colour similarity to rewarding model plants affects pollination in a food deceptive orchid, *Orchis boryi. Biological Journal of the Linnean Society*, *72*(3), 419–433.
- Gustafsson, S. (2000). Patterns of genetic variation in *Gymnadenia conopsea*, the fragrant orchid. *Molecular Ecology*, 9(11), 1863–1872.
- Hagerup, O. (1952). Bud autogamy in some northern orchids. *Phytomorphology*, 2, 51–61.
- Hamer, F. (1981). Las orquídeas de El Salvador, Tomo III. The Marie Selby Botanical Gardens.
- Hamer, F. (1982). Orchids of Nicaragua (Vols. 7-9). The Marie Selby Botanical Gardens.
- Hamer, F. (1984). Orchids of Nicaragua (Vols. 11-12). The Marie Selby Botanical Gardens.
- Hart, J. (1886). Self-fertilisation of Epidendrum variegatum. The Gardeners Chronicle.
- Hágsater, E. (1984). *Epidendrum trialatum*, una nueva especie centroamericana. *Epidendrum trialatum*, a new species from Central America. *Orquídea (Mexico City)*, 9(2), 355–360.
- Herring, T. (2007). A study of the reproductive systems of selected North American and Chinese Cypripedium species (Orchidaceae). [PhD Thesis]. Saint Louis University.
- Hogan, K. P. (1983). The pollination biology and breeding system of *Aplectrum hyemale* (Orchidaceae). *Canadian Journal of Botany*, *61*(7), 1906–1910.
- Holland, D. (1961). Native orchids of New Zealand. American Orchid Society Bulletin, 30, 372-373.
- Hollingsworth, P., Squirrell, J., Hollingsworth, M., Richards, A., Bateman, R., Bailey, J., & Ellis, R. (2006). Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis* (Orchidaceae). *Current Taxonomic Research on the British and European Flora*, 27–44.
- Householder, E., Janovec, J., Mozambite, A., Maceda, J., Wells, J., & Valega, R. (2010). Diversity, natural history, and conservation of *Vanilla* (Orchidaceae) in Amazonian wetlands of Madre de Dios, Peru. *Journal of the Botanical Research Institute of Texas*, 4(1), 227–243.
- Hu, A.-Q., Gale, S. W., Kumar, P., Saunders, R. M., Sun, M., & Fischer, G. A. (2017). Preponderance of clonality triggers loss of sex in *Bulbophyllum bicolor*, an obligately outcrossing epiphytic orchid. *Molecular Ecology*, 26(13), 3358–3372.
- Huda, M. K., & Wilcock, C. C. (2008). Impact of floral traits on the reproductive success of epiphytic and terrestrial tropical orchids. *Oecologia*, *154*(4), 731–741.

- Huda, M., & Wilcock, C. (2012). Rapid floral senescence following male function and breeding systems of some tropical orchids. *Plant Biology*, *14*(2), 278–284.
- Humaña, A. M., Cisternas, M. A., & Valdivia, C. E. (2008). Breeding system and pollination of selected orchids of the genus *Chloraea* (Orchidaceae) from central Chile. *Flora-Morphology, Distribution, Functional Ecology of Plants, 203*(6), 469–473.
- Iannotti, M. A., Stort, M., & Morales, M. (1987). O sistema reprodutivo de *Epidendrum rigidum* Jacq. (Orchidaceae). *Ciência E Cultura, 39*(7), 641–644.
- Ivri, Y., & Dafni, A. (1977). The pollination ecology of *Epipactis consimilis* Don (Orchidaceae) in Israel. New Phytologist, 79(1), 173–177.
- Jacquemyn, H., & Brys, R. (2015). Pollen limitation and the contribution of autonomous selfing to fruit and seed set in a rewarding orchid. *American Journal of Botany*, *102*(1), 67–72.
- Jacquemyn, H., Brys, R., Vandepitte, K., Honnay, O., & Roldan-Ruiz, I. (2006). Fine-scale genetic structure of life history stages in the food-deceptive orchid Orchis purpurea. Molecular Ecology, 15(10), 2801–2808.
- Jacquemyn, H., Micheneau, C., Roberts, D. L., & Pailler, T. (2005). Elevational gradients of species diversity, breeding system and floral traits of orchid species on Reunion Island. *Journal of Biogeog*raphy, 32(10), 1751–1761.
- Jaimes, I., & Ramírez, N. (1999). Breeding systems in a secondary deciduous forest in Venezuela: The importance of life form, habitat, and pollination specificity. *Plant Systematics and Evolution*, 215(1), 23–36.
- Janzen, D. H., DeVries, P., Gladstone, D. E., Higgins, M. L., & Lewinsohn, T. M. (1980). Self-and crosspollination of *Encyclia cordigera* (Orchidaceae) in Santa Rosa National Park, Costa Rica. *Biotropica*, 72–74.
- Jeanes, J. A. (2009). Resolution of the *Thelymitra variegata* (Orchidaceae) complex of southern Australia and New Zealand. Muelleria, 27(2), 149–170.
- Johansen, B. (1990). Incompatibility in *Dendrobium* (Orchidaceae). *Botanical Journal of the Linnean Society*, 103(2), 165–196.
- Johnson, S. (2000). Batesian mimicry in the non-rewarding orchid *Disa pulchra*, and its consequences for pollinator behaviour. *Biological Journal of the Linnean Society*, *71*(1), 119–132.
- Johnson, S., & Hobbhahn, N. (2010). Generalized pollination, floral scent chemistry, and a possible case of hybridization in the African orchid *Disa fragrans*. *South African Journal of Botany*, *76*(4), 739–748.
- Johnson, S., & Steiner, K. (1997). Long-tongued fly pollination and evolution of floral spur length in the *Disa draconis* complex (Orchidaceae). *Evolution*, *51*(1), 45–53.
- Johnson, S., Steiner, K., & Kurzweil, H. (1994). Auto-pollination in *Disa* species (Orchidaceae) of the Cape mountain cloud zone. *Lindleyana*, *9*, 3–6.
- Johnson, T. R., Stewart, S. L., Kauth, P., Kane, M. E., & Philman, N. (2009). Confronting assumptions about spontaneous autogamy in populations of *Eulophia alta* (Orchidaceae) in south Florida: Assessing the effect of pollination treatments on seed formation, seed germination and seedling development. *Botanical Journal of the Linnean Society*, 161(1), 78–88.
- Jones, D. (1972). The self-pollination of *Prasophyllum beaugleholei* W.H. Nicholls. *Victorian Naturalist*, *89*(5), 144–146.

- Jones, D. (1975). The pollination of Microtis parviflora R. Br. Annals of Botany, 39(3), 585-589.
- Jones, D. (1976). The self-pollination of *Phaius pictus* T.E. Hunt. Austral. Orchid Review, 210–213.
- Jones, D. (1985a). The pollination of *Gastrodia sesamoides* R. Br. In southern Victoria. *Victorian Naturalist*, 102, 52–54.
- Jones, D. (1985b). The self-pollination of *Epipogium roseum* (D. Don) Lindley. Orchadian, 8, 91–92.
- Jones, D., & Clements, M. (1987). Reinstatement of the genus *Cyrtostylis* R. Br. And its relationships with *Acianthus* R. Br. (Orchidaceae). *Lindleyana*, *2*, 156–160.
- Jones, D., & Gray, B. (1974). The pollination of *Calochilus campestris*. *American Orchid Society Bulletin*, 43, 604–606.

Jones, D., & Robinson, R. (1983). The pollination and occurrence of natural hybrids in Australian species of the genus *Dendrobium*. *Proceedings of the 8th Australian Orchid Conference*, 127–132.

- Jones, H. (1968). A note on *Schomburgkia* in Trinidad. *Lloydia*, *31*(3), 268.
- Jürgens, A., Bosch, S. R., Webber, A. C., Witt, T., Frame, D., & Gottsberger, G. (2009). Pollination biology of *Eulophia alta* (Orchidaceae) in Amazonia: Effects of pollinator composition on reproductive success in different populations. *Annals of Botany*, 104(5), 897–912.
- Kallunki, J. A. (1976). Population studies in *Goodyera* (Orchidaceae) with emphasis on the hybrid origin of *G. tesselata*. *Brittonia*, *28*(1), 53–75.
- Kallunki, J. A. (1981). Reproductive biology of mixed-species populations of *Goodyera* (Orchidaceae) in northern Michigan. *Brittonia*, *33*(2), 137–155.
- King, G. (1898). The orchids of the Sikkim Himalaya. *Annals of the Royal Botanic Garden, Culcutta*, 8, 1–342.
- Kipping, J. L. (1971). Pollination studies of native orchids [Master's Thesis]. San Francisco State College.
- Kirchner, O. von. (1922a). Über selbstbestäubung bei den orchideen. *Flora Oder Allgemeine Botanische Zeitung*, *115*(2-3), 103–129.
- Kirchner, O. von. (1922b). Zur Selbstbestäubung der Orchidaceen. Berichte Der Deutschen Botanischen Gesellschaft, 40, 317–321.
- Kjellsson, G., Rasmussen, F. N., & Dupuy, D. (1985). Pollination of *Dendrobium infundibulum, Cymbidium insigne* (Orchidaceae) and *Rhododendron lyi* (Ericaceae) by *Bombus eximius* (Apidae) in Thailand: A possible case of floral mimicry. *Journal of Tropical Ecology*, 289–302.
- Krahl, A. H., de Holanda, A. S., Krahl, D. R., Martucci, M. E., Gobbo-Neto, L., Webber, A. C., & Pansarin, E. R. (2019). Study of the reproductive biology of an Amazonian *Heterotaxis* (Orchidaceae) demonstrates the collection of resin-like material by stingless bees. *Plant Systematics and Evolution*, 305(4), 281–291.
- Krahl, A. H., Krahl, D. R., Valsko, J. J., Webber, A. C., & Pansarin, E. R. (2017). Evidence of reward production and pollination by *Centris* in *Encyclia* (Orchidaceae: Laeliinae): The reproductive biology of *Encyclia mapuerae*. Australian Journal of Botany, 65(3), 225–232.
- Kranzlin, F. (1903). Orchidearum generat et species: Vol. Chloraceae.
- Kropf, M. (2015). Platanthera hyperborea (Orchidaceae) only selfing? Rhodora, 117(972), 490–498.
- Lammi, A., Vallius, E., Vauhkonen, T., & Kuitunen, M. (2003). Outcrossing, hybridization, pollen quantity, and the evolution of deceptive pollination in *Dactylorhiza incarnata*. *Annales Botanici Fennici*, 331–338.

- Legras, S., Micheneau, C., Dupont, J., & Pailler, T. (2003). Biologie de la reproduction d'une espèce menacée de l'île de la Réunion (*Corymborkis corymbis*: Orchidaceae): Implications Pour Sa Conservation. *Revue d'écologie*.
- Lehnebach, C. A., & Robertson, A. W. (2004). Pollination ecology of four epiphytic orchids of New Zealand. *Annals of Botany*, *93*(6), 773–781.
- Lehnebach, C. A., Robertson, A. W., & Hedderley, D. (2005). Pollination studies of four New Zealand terrestrial orchids and the implication for their conservation. *New Zealand Journal of Botany*, 43(2), 467–477.
- Lehnebach, C., & Riveros, M. (2003). Pollination biology of the Chilean endemic orchid *Chloraea lamellata*. *Biodiversity & Conservation*, *12*(8), 1741–1751.
- Li, P., Luo, Y., Bernhardt, P., Yang, X., & Kou, Y. (2006). Deceptive pollination of the lady's slipper *Cypripedium tibeticum* (Orchidaceae). *Plant Systematics and Evolution*, *262*(1), 53-63.
- Li, X., Ding, X., Chu, B., Zhou, Q., Ding, G., & Gu, S. (2008). Genetic diversity analysis and conservation of the endangered Chinese endemic herb *Dendrobium officinale* Kimura et Migo (Orchidaceae) based on AFLP. *Genetica*, 133(2), 159–166.
- Lipow, S. R., Bernhardt, P., & Vance, N. (2002). Comparative rates of pollination and fruit set in widely separated populations of a rare orchid (*Cypripedium fasciculatum*). *International Journal of Plant Sciences*, 163(5), 775–782.
- Little, K. J., Dieringer, G., & Romano, M. (2005). Pollination ecology, genetic diversity and selection on nectar spur length in *Platanthera lacera* (Orchidaceae). *Plant Species Biology*, *20*(3), 183–190.
- Liu, K.-W., Liu, Z.-J., Huang, L., Li, L.-Q., Chen, L.-J., & Tang, G.-D. (2006). Self-fertilization strategy in an orchid. *Nature*, *441*(7096), 945–946.
- Liu, Z.-J., Chen, L.-J., Liu, K.-W., Li, L.-Q., Rao, W.-H., Zhang, Y.-T., Tang, G.-D., & Huang, L.-Q. (2013). Adding perches for cross-pollination ensures the reproduction of a self-incompatible orchid. *PLoS One*, 8(1), e53695.
- Lloyd, D., & Schoen, D. (1992). Self-fertilization and cross-fertilization in plants. I. Functional dimensions. *International Journal of Plant Sciences*, *153*(3), 358–369.
- Luer, C. A. (1972). The native orchids of Florida. New York Botanical Garden.
- Luer, C. A. (1977). Icones Pleurothallidinarum (Orchidaceae): Miscellaneous species in the Pleurothallidinae. *Selbyana*, *3*(3/4), 203–412.
- Luer, C. A. (1978). *Dracula*, a new genus in the Pleurothallidinae. *Selbyana*, *2*(2/3), 190–198.
- Luyt, R. P. (2002). *Pollination and evolution of the genus* Mystacidium (*Orchidaceae*) [Master's Thesis]. University of Natal, School of Botany; Zoology.
- Maad, J., & Reinhammar, L. G. (2004). Incidence of geitonogamy differs between two populations in the hawkmoth-pollinated *Platanthera bifolia* (Orchidaceae). *Canadian Journal of Botany*, 82(11), 1586–1593.
- Mallet, B., Martos, F., Blambert, L., Pailler, T., & Humeau, L. (2014). Evidence for isolation-by-habitat among populations of an epiphytic orchid species on a small oceanic island. *PloS One*, *9*(2), e87469.
- Martel, C., Cairampoma, L., Stauffer, F. W., & Ayasse, M. (2016). *Telipogon peruvianus* (Orchidaceae) flowers elicit pre-mating behaviour in *Eudejeania* (Tachinidae) males for pollination. *PLoS One*, *11*(11), e0165896.

- Martens, P. (1926). L'Autogamie chez l'Orchis et chez quelques autres orchidées. Bulletin de La Société Royale de Botanique de Belgique/Bulletin van de Koninklijke Belgische Botanische Vereniging, 59(Fasc. 1), 69–88.
- Matsuda, Y., & Sugiura, N. (2019). Specialized pollination by honeybees in *Cymbidium dayanum*, a fallWinter flowering orchid. *Plant Species Biology*, *34*(1), 19–26.
- Matsui, K., Ushimaru, T., & Fujita, N. (2001). Pollinator limitaton in a deceptive orchid, *Pogonia japonica*, on a floating peat mat. *Plant Species Biology*, *16*(3), 231–235.
- McMullen, C. K. (1987). Breeding systems of selected Galapagos Islands angiosperms. American Journal of Botany, 74(11), 1694–1705.
- Medley, M. E. (1979). Some aspects of the life history of Triphora trianthophora (Sw.) Rydb (three birds orchid) with special reference to its pollination [Doctoral Dissertation]. Andrews University, School of Graduate Studies.
- Meekers, T., Hutchings, M. J., Honnay, O., & Jacquemyn, H. (2012). Biological flora of the British Isles: *Gymnadenia conopsea* s.l. *Journal of Ecology*, *100*(5), 1269–1288.
- Mehrhoff III, L. A. (1983). Pollination in the genus *Isotria* (Orchidaceae). *American Journal of Botany*, 70(10), 1444–1453.
- Meléndez-Ackerman, E. J., & Ackerman, J. D. (2001). Density-dependent variation in reproductive success in a terrestrial orchid. *Plant Systematics and Evolution*, 227(1), 27–36.
- Micheneau, C., Fournel, J., Gauvin-Bialecki, A., & Pailler, T. (2008). Auto-pollination in a long-spurred endemic orchid (*Jumellea stenophylla*) on Reunion Island (Mascarene Archipelago, Indian Ocean). *Plant Systematics and Evolution*, 272(1), 11–22.
- Micheneau, C., Fournel, J., & Pailler, T. (2006). Bird pollination in an angraecoid orchid on Reunion Island (Mascarene Archipelago, Indian Ocean). *Annals of Botany*, *97*(6), 965–974.
- Micheneau, C., Fournel, J., Warren, B. H., Hugel, S., Gauvin-Bialecki, A., Pailler, T., Strasberg, D., & Chase, M. W. (2010). Orthoptera, a new order of pollinator. *Annals of Botany*, *105*(3), 355–364.
- Millner, H. J., McCrea, A. R., & Baldwin, T. C. (2015). An investigation of self-incompatibility within the genus *Restrepia*. *American Journal of Botany*, *102*(3), 487–494.
- Moggridge, J. T. (1865). Observations on some orchids of the South of France. *Botanical Journal of the Linnean Society*, 8(32), 256–258.
- Montalvo, A. M., & Ackerman, J. D. (1987). Limitations to fruit production in *Ionopsis utricularioides* (Orchidaceae). *Biotropica*, 24–31.
- Moore, S. (1876). Mascarene orchidology. Journal of Botany, 289–292.
- Moreira, G. R., Correa, C. A., & Mugrabi-Oliveira, E. (1996). Pollination of *Habenaria pleiophylla* Hoehne & Schlechter (Orchidaceae) by *Heliconius erato phyllis* Fabricius (Lepidoptera, Nymphalidae). *Revista Brasileira de Zoologia*, *13*(3), 791–798.
- Nannfeldt, J. (1946). Tre för Norden nya *Epipactis*-arter, *E. persica* Hsausskn., *E. leptochila* (Godf.) Godf., Och *E. purpurata* Sm. *Botaniska Notiser*, 1–28.
- Nargar, K., Molina, S., Wagner, N., Nauheimer, L., Micheneau, C., & Clements, M. A. (2019). Australasian orchid diversification in time and space: Molecular phylogenetic insights from the beard orchids (*Calochilus*, Diurideae). *Australian Systematic Botany*, 31(6), 389–408.
- Neiland, M. R. M., & Wilcock, C. C. (1995). Maximisation of reproductive success by European Orchidaceae under conditions of infrequent pollination. *Protoplasma*, *187*(1-4), 39–48.

- Neiland, M. R. M., & Wilcock, C. C. (1998). Fruit set, nectar reward, and rarity in the Orchidaceae. *American Journal of Botany*, 85(12), 1657–1671.
- Newstrom, L., & Robertson, A. (2005). Progress in understanding pollination systems in New Zealand. New Zealand Journal of Botany, 43(1), 1–59.
- Nilsson, L. (1980). The pollination ecology of *Dactylorhiza sambucina* (Orchidaceae). *Botaniska Notiser*, 133, 367–385.
- Nilsson, L. A. (1981). The pollination ecology of *Listera ovata* (Orchidaceae). *Nordic Journal of Botany*, *1*(4), 461–480.
- Nilsson, L. A. (1983a). Anthecology of Orchis mascula (Orchidaceae). Nordic Journal of Botany, 3(2), 157–179.
- Nilsson, L. A. (1983b). Processes of isolation and introgressive interplay between *Platanthera bifolia* (L.) Rich and *P. chlorantha* (Custer) Reichb. (Orchidaceae). *Botanical Journal of the Linnean Society*, 87(4), 325–350.
- Nilsson, L. A. (1984). Anthecology of *Orchis morio* (Orchidaceae) at its outpost in the north. *Nova Acta Regiae Societatis Scientarium Upsaliensis*, *3*, 167–179.
- Nilsson, L. A., & Rabakonandrianina, E. (1988). Hawk-moth scale analysis and pollination specialization in the epilithic Malagasy endemic *Aerangis ellisii* (Reichenb. Fil.) Schltr. (Orchidaceae). *Botanical Journal of the Linnean Society*, 97(1), 49–61.
- Niu, S.-C., Huang, J., Xu, Q., Li, P.-X., Yang, H.-J., Zhang, Y.-Q., Zhang, G.-Q., Chen, L.-J., Niu, Y.-X., Luo, Y.-B., & others. (2018). Morphological type identification of self-incompatibility in *Dendrobium* and its phylogenetic evolution pattern. *International Journal of Molecular Sciences*, 19(9), 2595.
- Niu, S.-C., Huang, J., Zhang, Y.-Q., Li, P.-X., Zhang, G.-Q., Xu, Q., Chen, L.-J., Wang, J.-Y., Luo, Y.-B., & Liu, Z.-J. (2017). Lack of S-RNase-based gametophytic self-incompatibility in orchids suggests that this system evolved after the monocot-eudicot split. *Frontiers in Plant Science*, *8*, 1106.
- Nunes, C. E., Amorim, F. W., Mayer, J. L. S., & Sazima, M. (2016). Pollination ecology of two species of *Elleanthus* (Orchidaceae): Novel mechanisms and underlying adaptations to hummingbird pollination. *Plant Biology*, 18(1), 15–25.
- Oakwood, M. (1990). *Population genetics and pollination ecology of* Chiloglottis trapeziformis Fitzg. (Orchidaceae) [Doctoral Dissertation]. Macquarie University.
- Oh, G. S., Chung, M. Y., Chung, S. G., & Chung, M. G. (2001). Contrasting breeding systems: *Liparis kumokiri* and *L. makinoana* (Orchidaceae). *Annales Botanici Fennici*, 281–284.
- Ortiz-Barney, E., & Ackerman, J. D. (1999). The cost of selfing in *Encyclia cochleata (Orchidaceae)*. Plant Systematics and Evolution, 219(1), 55–64.
- Ospina-Calderón, N. H., Duque-Buitrago, C. A., Tremblay, R. L., & Tupac Otero, J. (2015). Pollination ecology of *Rodriguezia granadensis* (Orchidaceae). *Lankesteriana*, *15*(2), 129–139.
- Pais, M. (1969). Morphologie, structure et cytochemie de la surface stigmatique d'*Ophrys lutea* L (Orchidee). *Portugaliae Acta Biologica*, *11*, 339–346.
- Pansarin, E., Alves-dos-Santos, I., & Pansarin, L. (2017). Comparative reproductive biology and pollinator specificity among sympatric *Gomesa* (Orchidaceae: Oncidiinae). *Plant Biology*, 19(2), 147–155.
- Pansarin, E., & Pansarin, L. (2011). Reproductive biology of *Trichocentrum pumilum*: An orchid pollinated by oil-collecting bees. *Plant Biology*, 13(4), 576–581.

- Pansarin, E., Pansarin, L., Martucci, M., & Gobbo-Neto, L. (2016). Self-compatibility and specialisation in a fly-pollinated *Acianthera* (Orchidaceae: Pleurothallidiinae). *Australian Journal of Botany*, 64(4), 359–367.
- Pansarin, E. R. (2008). Reproductive biology and pollination of *Govenia utriculata*: A syrphid fly orchid pollinated through a pollen-deceptive mechanism. *Plant Species Biology*, *23*(2), 90–96.
- Pansarin, E. R., & Amaral, M. do C. E. do. (2008). Reproductive biology and pollination mechanisms of *Epidendrum secundum* (Orchidaceae). Floral variation: A consequence of natural hybridization? *Plant Biology*, 10(2), 211–219.
- Pansarin, E. R., & do Amaral, M. do C. E. (2008). Pollen and nectar as a reward in the basal epidendroid *Psilochilus modestus* (Orchidaceae: Triphoreae): A Study of Floral Morphology, Reproductive Biology and Pollination Strategy. *Flora-Morphology, Distribution, Functional Ecology of Plants, 203*(6), 474–483.
- Pansarin, E. R., & do Amaral, M. do C. E. (2009). Reproductive biology and pollination of southeastern Brazilian Stanhopea Frost ex Hook. (Orchidaceae). Flora-Morphology, Distribution, Functional Ecology of Plants, 204(3), 238–249.
- Pansarin, E. R., & Pansarin, L. M. (2011). The family Orchidaceae in the Serra do Japi, São Paulo State, Brazil. Springer Science & Business Media.
- Pansarin, E. R., & Pansarin, L. M. (2017). Crane flies and microlepidoptera also function as pollinators in *Epidendrum* (Orchidaceae: Laeliinae): The reproductive biology of *E. avicula*. *Plant Species Biology*, 32(3), 200–209.
- Pansarin, E. R., Pansarin, L. M., & Alves-dos-Santos, I. (2015). Floral features, pollination biology, and breeding system of *Comparettia coccinea* (Orchidaceae: Oncidiinae). *Flora-Morphology, Distribution, Functional Ecology of Plants, 217*, 57–63.
- Pansarin, L. M., Pansarin, E. R., & Sazima, M. (2008). Facultative autogamy in *Cyrtopodium polyphyllum* (Orchidaceae) through a rain-assisted pollination mechanism. *Australian Journal of Botany*, 56(4), 363–367.
- Parra-Tabla, V., Vargas, C. F., Magaña-Rueda, S., & Navarro, J. (2000). Female and male pollination success of *Oncidium ascendens* Lindey (Orchidaceae) in two contrasting habitat patches: Forest vs agricultural field. *Biological Conservation*, 94(3), 335–340.
- Pauw, A. (2006). Floral syndromes accurately predict pollination by a specialized oil-collecting bee (*Rediviva peringueyi*, Melittidae) in a guild of South African orchids (Coryciinae). *American Journal* of Botany, 93(6), 917–926.
- Peakall, R. (1989). The unique pollination of *Leporella fimbriata* (Orchidaceae): Pollination by pseudocopulating male ants (*Myrmecia urens*, Formicidae). *Plant Systematics and Evolution*, 167(3), 137– 148.
- Peakall, R., & Beattie, A. J. (1991). The genetic consequences of worker ant pollination in a self-compatible, clonal orchid. *Evolution*, 45(8), 1837–1848.
- Peakall, R., & Beattie, A. J. (1996). Ecological and genetic consequences of pollination by sexual deception in the orchid *Caladenia tentactulata*. *Evolution*, *50*(6), 2207–2220.
- Pederson, H. (1995). Anthecological observations on *Dendrobium longibrachteatum*: A species pollinated by facultatively anthophilous insects. *Lindleyana*, *10*(1), 19–28.

- Pedron, M., Buzatto, C. R., Singer, R. B., Batista, J. A., & Moser, A. (2012). Pollination biology of four sympatric species of *Habenaria* (Orchidaceae: Orchidinae) from southern Brazil. *Botanical Journal* of the Linnean Society, 170(2), 141–156.
- Pellegrino, G., Bellusci, F., & Musacchio, A. (2010). The effects of inflorescence size and flower position on female reproductive success in three deceptive orchids. *Botanical Studies*, *51*(3).
- Pellegrino, G., Bellusci, F., & Palermo, A. M. (2015). Effects of population structure on pollen flow, clonality rates and reproductive success in fragmented *Serapias lingua* populations. *BMC Plant Biology*, 15(1), 1–10.
- Peter, C. I. (2009). *Pollination, floral deception and evolutionary processes in* Eulophia (*Orchidaceae*) and *its allies* [Doctoral Dissertation]. University of KwaZulu-Natal, Pietermaritzburg.
- Peter, C. I., & Johnson, S. D. (2006). Anther cap retention prevents self-pollination by elaterid beetles in the South African orchid *Eulophia foliosa*. *Annals of Botany*, *97*(3), 345–355.
- Peter, C., & Johnson, S. (2009a). Autonomous self-pollination and pseudo-fruit set in South African species of *Eulophia* (Orchidaceae). *South African Journal of Botany*, *75*(4), 791–797.
- Peter, C., & Johnson, S. (2009b). Pollination by flower chafer beetles in *Eulophia ensata* and *Eulophia welwitschii* (Orchidaceae). *South African Journal of Botany*, *75*(4), 762–770.
- Pinheiro, F., Cafasso, D., Cozzolino, S., & Scopece, G. (2015). Transitions between self-compatibility and self-incompatibility and the evolution of reproductive isolation in the large and diverse tropical genus *Dendrobium* (Orchidaceae). *Annals of Botany*, 116(3), 457–467.
- Ramirez, N., & Seres, A. (1994). Plant reproductive biology of herbaceous monocots in a Venezuelan tropical cloud forest. *Plant Systematics and Evolution*, *190*(3), 129–142.
- Razanajatovo, M., Maurel, N., Dawson, W., Essl, F., Kreft, H., Pergl, J., Pyšek, P., Weigelt, P., Winter, M., & Van Kleunen, M. (2016). Plants capable of selfing are more likely to become naturalized. *Nature Communications*, 7(1), 1–9.
- Reiche, K. (1910). Orchidaceae Chilensis. Anales Del Museo Nacional de Chile, Sección 2, Botánica, Entrega, 18, 1–88.
- Reichenbach, H. (1877). Orchideae Roezlianae novae seu criticae. Linnaea, 41, 1–98.
- Reinhard, H. (1977). Autogamie bei europäischen Orchideen. Orchidee, 28, 178-182.
- Ribeiro, M., Köhler, A., & Boelter, C. (2006). Polinização de *Acianthera aphtosa* (Lindl.) Pridgeon & MW Chase (Orchidaceae) por Otitidae (Diptera). *Uruguaiana*, *13*(2), 85-89.
- Rico Gray, V., & Thien, L. (1987). Some aspects of the reproductive biology of *Schomburgkia tibicinis* Batem (Orchidaceae) in Yucatan, Mexico. *Brenesia*, *28*, 13–24.
- Ridley, H. N. (1885). The orchids of Madagascar. *Botanical Journal of the Linnean Society*, *21*(137), 456–522.
- Ridley, H. N. (1888). A revision of the genera *Microstylis* and *Malaxis*. *Botanical Journal of the Linnean Society*, *24*(162), 308–351.
- Roberts, D. L. (2003). Pollination biology: The role of sexual reproduction in orchid conservation. In K. Dixon, S. Kell, R. Barrett, & P. Cribb (Eds.), *Orchid Conservation* (pp. 113–136). Natural History Publications.
- Robertson, J. L., & Wyatt, R. (1990). Evidence for pollination ecotypes in the yellow-fringed orchid, *Platanthera ciliaris. Evolution*, 44(1), 121–133.

- Rodolphe, G., Séverine, B., Michel, G., & Pascale, B. (2011). Biodiversity and evolution in the *Vanilla* genus. In *The Dynamical Processes of Biodiversity Case studies of evolution and spatial distribution* (pp. 1–27).
- Rodríguez-Robles, J. A., Meléndez, E. J., & Ackerman, J. D. (1992). Effects of display size, flowering phenology, and nectar availability on effective visitation frequency in *Comparettia falcata* (Orchidaceae). *American Journal of Botany*, *79*(9), 1009–1017.
- Rogers, R. (1913). Mechanisms of pollination in certain Australian orchids. *Transactions and Proceedings* of the Royal Society of South Australia, 37, 48–61.
- Rogers, R. S. (1921). Notes on the Gynostemium in the genus *Diuris*, and on the pollinary mechanism in *Phajus*. *Transactions and Proceedings of the Royal Society of South Australia*, 45, 264–269.
- Romand-Monnier, F. (2014). *Habenaria parviflora* Lindl. In *IUCN*. https://eol.org/pages/1113156
- Salguero-Faría, J. A., & Ackerman, J. D. (1999). A nectar reward: Is more better? *Biotropica*, 31(2), 303-311.
- Sanguinetti, A., Buzatto, C. R., Pedron, M., Davies, K. L., Ferreira, P. M. de A., Maldonado, S., & Singer, R. B. (2012). Floral features, pollination biology and breeding system of *Chloraea membranacea* Lindl. (Orchidaceae: Chloraeinae). *Annals of Botany*, 110(8), 1607–1621.
- Sather, N. (Ed.). (1996). Appendix N: Recovery strategy for the western prairie fringed orchid on the Sheyenne National Grassland. In Western Prairie Fringed Orchid Recovery Plan (Platanthera praeclara). US Fish and Wildlife Service.
- Sauleda, R., Wunderlin, R., & Hansen, B. (1984). Transfers in Florida orchids. Phytologia.
- Sazima, M. (1978). Polinização por moscas em *Bulbophyllum warmingianum* Cogn. (Orchidaceae), na Serra do Cipó, Minas Gerais, Brasil. *Revista Brasileira de Botânica*, *1*, 133–138.
- São Leão, L. C., de Sá-Haiad, B., de Araujo Rodarte, A. T., Pimentel, R. R., Benevides, C. R., de Santiago-Fernandes, L. D. R., & de Lima, H. A. (2019). Reproductive biology of two synchronopatric neotropical species of *Epidendrum* (Orchidaceae). *Flora*, 251, 95–104.
- Schelpe, E. (1970). Self-pollination, cleistogamy and apomixis among Southern African orchids. *South African Orchid Journal*, 9–10.
- Schemske, D. W. (1980). Evolution of floral display in the orchid Brassavola nodosa. Evolution, 489–493.
- Schlechter, F. (1914). Gruppe 65: Sarcanthinae, untergruppe 2: Vandeae. Die Orchideen, III, 555-611.
- Schlechter, R. (1911). Die polychondreae (Neottiinae Pfitz.) Und ihre systematische einteilung. *Botanische Jahrbücher Für Systematik, Pflanzengeschichte, Und Pflanzengeographie, 45*, 375–410.
- Schlechter, R. (1922). Die Orchideenflora der sudamerikanische Kordillerenstaaten. V. Bolivia. In F. Fedde (Ed.), Repertorium Specierum Novarum Regni Vegetabilis. Centralblatt für Sammlung und Veroffentlichung von Einzeldiagnosen neuer Pflanzen (Vol. 10, pp. 1–80).
- Scott, J. (1864). On the individual sterility and cross-impregnation of certain species of *Oncidium*. *Botanical Journal of the Linnean Society*, *8*(31), 162–167.
- Sharp, P. C. (2004). Platanthera ciliaris (*L.*) *Lindl.: Yellow-fringed orchis*. Conservation and Research Plan for New England.
- Shen, X., Liu, C., & Pan, K. (2014). Reproductive biological characteristics of *Dendrobium* species. *Reproductive Biology of Plants*, 195.

- Silva-Pereira, V., de Camargo Smidt, E., & Borba, E. L. (2007). Isolation mechanisms between two sympatric Sophronitis (Orchidaceae) species endemic to Northeastern Brazil. Plant Systematics and Evolution, 269(3), 171–182.
- Singer, R. B. (2001). Pollination biology of *Hebenaria parviflora* (Orchidaceae: Habenariinae) in southeastern Brazil. *Darwiniana*, 201–207.
- Singer, R. B. (2002a). The pollination biology of *Sauroglossum elatum* Lindl. (Orchidaceae: Spiranthinae): Moth-pollination and protandry in neotropical Spiranthinae. *Botanical Journal of the Linnean Society*, 138(1), 9–16.
- Singer, R. B. (2002b). The pollination mechanism in *Trigonidium obtusum* Lindl. (Orchidaceae: Maxillariinae): Sexual mimicry and trap-flowers. *Annals of Botany*, *89*(2), 157–163.
- Singer, R. B., Breier, T. B., Flach, A., & Farias-Singer, R. (2007). The pollination mechanism of Habenaria pleiophylla Hoehne & Schlechter (Orchidaceae: Orchidinae). Functional Ecosystems and Communities, 1(1), 10–14.
- Singer, R. B., Gravendeel, B., Cross, H., & Ramirez, S. R. (2008). The use of orchid pollinia or pollinaria for taxonomic identification. *Selbyana*, 6–19.
- Singer, R. B., & Sazima, M. (1999). The pollination mechanism in the *"Pelexia* alliance" (Orchidaceae: Spiranthinae). *Botanical Journal of the Linnean Society*, *131*(3), 249–262.
- Singer, R. B., & Sazima, M. (2001a). Flower morphology and pollination mechanism in three sympatric Goodyerinae orchids from southeastern Brazil. *Annals of Botany*, *88*(6), 989–997.
- Singer, R. B., & Sazima, M. (2001b). The pollination mechanism of three sympatric *Prescottia* (Orchidaceae: Prescottinae) species in southeastern Brazil. *Annals of Botany*, *88*(6), 999–1005.
- Sletvold, N., & Ågren, J. (2010). Pollinator-mediated selection on floral display and spur length in the orchid *Gymnadenia conopsea*. International Journal of Plant Sciences, 171(9), 999–1009.
- Smidt, E. C., Silva-Pereira, V., & Borba, E. L. (2006). Reproductive biology of two Cattleya (Orchidaceae) species endemic to north-eastern Brazil. Plant Species Biology, 21(2), 85–91.
- Smith, G. R., & Snow, G. E. (1976). Pollination ecology of *Platanthera (Habenaria) ciliaris* and *P. blephar-iglottis* (Orchidaceae). *Botanical Gazette*, 137(2), 133–140.
- Smith, J. J. (1905). Die Orchideen von Ambon. Landsdrukkerij.
- Smith, J. J. (1907a). Die Orchideen von Java. EJ Brill.
- Smith, J. J. (1907b). Die Orchideen von Nova Guinea. Bulletin Du Département de L'Agriculture Aux Indes Néerlandaises, 13.
- Smith, J. J. (1928). Zelfbevruchting bij Orchideeen. Natuurkundig Tijdschrift Voor Nederlandsch Indië, 88, 122–140.
- Snow, A. A., & Whigham, D. F. (1989). Costs of flower and fruit production in *Tipularia discolor* (Orchidaceae). *Ecology*, *70*(5), 1286–1293.
- Son, H.-D., Im, H.-T., & Choi, S.-W. (2019). Pollination of *Cleisostoma scolopendrifolium* (Orchidaceae) by megachilid bees and determinants of fruit set in southern South Korea. *Journal of Ecology and Environment*, 43(1), 1–5.
- Sosa, V., & Rodríguez-Angulo, M. (2000). Pollination and natural history of *Mormodes tuxtlensis* (Catasetinae: Orchidaceae). *Lindleyana*, *15*(1), 53–58.

- Srimuang, K.-o., Watthana, S., Rangsayatorn, N., Eungwanichayapant, P. D., & others. (2010). Aspects of biosubsistence in *Sirindhornia* (Orchidaceae): Are the narrow endemics more reproductively restricted than their widespread relative? *Annales Botanici Fennici*, 47, 449–459.
- Steiner, K. E. (1998). The evolution of beetle pollination in a South African orchid. American Journal of Botany, 85(8), 1180–1193.
- Steiner, K. E., Whitehead, V., & Johnson, S. (1994). Floral and pollinator divergence in two sexually deceptive South African orchids. *American Journal of Botany*, 81(2), 185–194.
- Stort, M., & Martins, P. (1980). Autopolinização e polinização cruzada em algumas espécies do gênero Cattleya (Orchidaceae). Ciência E Cultura, 32, 1080–1084.
- Stort, M. N. S., & dos Santos Pavanelli, E. A. (1986). Crossing systems in *Epidendrum nocturnum* Jacq. (Orchidaceae). *Revista de Biología Tropical*, 34(1), 59-62.
- Stort, M. N. S., & Galdino, G. (1984). Self-and cross-pollination in some species of the genus *Laelia* Lindl. (Orchidaceae). *Revista Brasileira de Genética*, 7(4).
- Stoutamire, W. (1975). Pseudocopulation in Australian terrestrial orchids. *American Orchid Society Bulletin*.
- Stoutamire, W. (1983). Early growth in North American terrestrial orchid seedlings. North American Terrestrial Orchids Symposium II: Proceedings and Lectures, 14–24.
- Stoutamire, W. P. (1974). Relationships of the purple-fringed orchids *Platanthera psycodes* and *P. grandiflora*. *Brittonia*, 26(1), 42–58.
- Suetsugu, K. (2015a). Autonomous self-pollination and insect visitors in partially and fully mycoheterotrophic species of *Cymbidium* (Orchidaceae). *Journal of Plant Research*, *128*(1), 115–125.
- Suetsugu, K. (2015b). Autonomous self-pollination in the nectarless orchid *Pogonia minor*. *Plant Species Biology*, *30*(1), 37–41.
- Suetsugu, K. (2019). Rain-triggered self-pollination in *Liparis kumokiri*, an orchid that blooms during the rainy season. *Ecology*, *100*(7), e02683.
- Sugiura, N., Fujie, T., Inoue, K., & Kitamura, K. (2001). Flowering phenology, pollination, and fruit set of *Cypripedium macranthos var. rebunense*, a threatened lady's slipper (Orchidaceae). *Journal of Plant Research*, 114(2), 171–178.
- Tan, K. H., & Tan, L. T. (2019). Movements of floral parts and roles of the tooth on the column wall of Bulbophyllum praetervisum (Orchidaceae) flower in pollination by Dacini fruit flies (Diptera: Tephritidae). Journal of Pollination Ecology, 24.
- Tashev, A., Vitkova, A., & Russakova, V. (2006). Distribution of *Ophrys apifera* Huds. (Orchidaceae) in Bulgaria. *Flora Mediterranea*, *16*, 247–252.
- Tasmanian threatened species listing statement: Corunastylis nuda. (2017). Department of Primary Industries, Parks, Water and Environment. https://www.threatenedspecieslink.tas.gov .au/Pages/Corunastylis-nuda.aspx
- Tałałaj, I., & Brzosko, E. (2008). Selfing potential in *Epipactis palustris, E. helleborine* and *E. atrorubens* (Orchidaceae). *Plant Systematics and Evolution, 276*(1), 21–29.
- Tałałaj, I., Kotowicz, J., Brzosko, E., Ostrowiecka, B., Aleksandrowicz, O., & Wróblewska, A. (2019). Spontaneous caudicle reconfiguration in *Dactylorhiza*. *Plant Systematics and Evolution*, 305(4), 269–280.

- Thakur, U., & Dongarwar, N. (2019). Species specific pollination and seed germination in *Rhynchostylis* retusa Blume (Orchidaceae). Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences, 5(2), 387–395.
- Thien, L. (1973). Isolating mechanisms in the genus Calopogon. American Orchid Society Bulletin.
- Thien, L. B. (1969). Mosquito pollination of *Habenaria obtusata* (Orchidaceae). *American Journal of Botany*, *56*(2), 232–237.
- Thien, L. B., & Marcks, B. G. (1972). The floral biology of *Arethusa bulbosa*, *Calopogon tuberosus*, and *Pogonia ophioglossoides* (Orchidaceae). *Canadian Journal of Botany*, *50*(11), 2319–2325.
- Thompson, G. (1879). On the means of fertilization among some New Zealand orchids. *Transactions* of the Proceedings of the New Zealand Institute, 11, 418–426.
- Torretta, J. P., Gomiz, N. E., Aliscioni, S. S., & Bello, M. E. (2011). Biología reproductiva de *Gomesa bifolia* (Orchidaceae, Cymbidieae, Oncidiinae). *Darwiniana*, 16–24.
- Travers, S. E., Anderson, K., Vitt, P., & Harris, M. O. (2018). Breeding system and inbreeding depression in the rare orchid, *Platanthera praeclara*, in a fragmented grassland landscape. *Botany-Botanique*, 96(3), 151–159.
- Tremblay, R. L., Ackerman, J. D., Zimmerman, J. K., & Calvo, R. N. (2005). Variation in sexual reproduction in orchids and its evolutionary consequences: A spasmodic journey to diversification. *Biological Journal of the Linnean Society*, 84(1), 1–54.
- Tremblay, R. L., Pomales-Hernández, G., & Méndez-Cintrón, M. de L. (2006). Flower phenology and sexual maturation: Partial protandrous behavior in three species of orchids. *Caribbean Journal of Science*, 42(1), 75–80.
- Ueno, S., Rodrigues, J. F., Alves-Pereira, A., Pansarin, E. R., & Veasey, E. A. (2015). Genetic variability within and among populations of an invasive, exotic orchid. *AoB Plants*, *7*.
- Valdivia, C. E., Cisternas, M. A., & Verdugo, G. S. (2010). Reproductive biology aspects of two species of the genus *Gavilea* (Orchidaceae, Chloraeinae) in populations from Central Chile. *Gayana Botanica*, 67(1), 44–51.
- Vale, A., Navarro, L., Rojas, D., & Alvarez, J. C. (2011). Breeding system and pollination by mimicry of the orchid *Tolumnia guibertiana* in Western Cuba. *Plant Species Biology*, *26*(2), 163–173.
- Vale, A., Rojas, D., Alvarez, J., & Navarro, L. (2011). Breeding system and factors limiting fruit production in the nectarless orchid *Broughtonia lindenii*. *Plant Biology*, 13, 51–61.
- Vallius, E. (2000). Position-dependent reproductive success of flowers in *Dactylorhiza maculata* (Orchidaceae). *Functional Ecology*, 14, 573–579.
- Van Dam, A. R., Householder, J. E., & Lubinsky, P. (2010). Vanilla bicolor Lindl. (Orchidaceae) from the Peruvian Amazon: Auto-fertilization in Vanilla and notes on floral phenology. Genetic Resources and Crop Evolution, 57(4), 473–480.
- van der Niet, T., Hansen, D. M., & Johnson, S. D. (2011). Carrion mimicry in a South African orchid: Flowers attract a narrow subset of the fly assemblage on animal carcasses. *Annals of Botany*, *107*(6), 981–992.
- Vega, Y., & Marques, I. (2015). Both biotic and abiotic factors influence floral longevity in three species of *Epidendrum* (Orchidaceae). *Plant Species Biology*, *30*(3), 184–192.

- Vieira, B., Pansarin, L., Martucci, M., Gobbo-Neto, L., & Pansarin, E. (2018). Pollinarium size as a hybridisation barrier between sympatric inter-compatible orchids. *Australian Journal of Botany*, 65(7), 497–506.
- Vieira, M. F., Andrade, M. R. S., Bittencourt, N. S., & de Carvalho-Okano, R. M. (2007). Flowering phenology, nectary structure and breeding system in *Corymborkis flava* (Spiranthoideae: Tropidieae), a terrestrial orchid from a Neotropical forest. *Australian Journal of Botany*, 55(6), 635–642.
- Vitt, P., & Campbell, C. S. (1997). Reproductive biology of *Isotria medeoloides* (Orchidaceae). *Rhodora*, 56–63.
- Vogel, S. (1984). The *Diascia* flower and its beeand-oil-based symbiosis in southern Africa. *Acta Botan-ica Neerlandica*, 33(4), 509–518.
- Wallace, L. E. (2002a). Biological investigations in the genus Platanthera (Orchidaceae): Conservation issues in Platanthera leucophaea and evolutionary diversification in section Limnorchis [Doctoral Dissertation]. The Ohio State University.
- Wallace, L. E. (2002b). Examining the effects of fragmentation on genetic variation in *Platanthera leucophaea* (Orchidaceae): Inferences from allozyme and random amplified polymorphic DNA markers. *Plant Species Biology*, 17(1), 37–49.
- Wallace, L. E. (2003). The cost of inbreeding in *Platanthera leucophaea* (Orchidaceae). *American Journal* of Botany, 90(2), 235–242.
- Wheelwright, N. T., Dukeshire, E. E., Fontaine, J. B., Gutow, S. H., Moeller, D. A., Schuetz, J. G., Smith, T. M., Rodgers, S. L., & Zink, A. G. (2006). Pollinator limitation, autogamy and minimal inbreeding depression in insect-pollinated plants on a Boreal Island. *The American Midland Naturalist*, 155(1), 19–38.
- Whigham, D. F., & McWethy, M. (1980). Studies on the pollination ecology of *Tipularia discolor* (Orchidaceae). *American Journal of Botany*, 67(4), 550–555.
- Whigham, D. F., & O'Neill, J. P. (1991). Dynamics of flowering and fruit production in two eastern North American terrestrial orchids, *Tipularia discolor* and *Liparis lilifolia*. In *Population Ecology of Terrestrial Orchids* (pp. 89–101). SPB Academic Publishing.
- Whitehead, M., Linde, C., & Peakall, R. (2015). Pollination by sexual deception promotes outcrossing and mate diversity in self-compatible clonal orchids. *Journal of Evolutionary Biology*, 28(8), 1526– 1541.
- Wilfret, G. J. (1968). *Genome and Karyote Relationships in the Genus* Dendrobium [Doctoral Dissertation]. University of Hawaii.
- Willems, J., & Lahtinen, M.-L. (1997). Impact of pollination and resource limitation on seed production in a border population of *Spiranthes spiralis* (Orchidaceae). *Acta Botanica Neerlandica*, 46(4), 365– 375.
- Williamson, G. (1984). Observations of a mechanism by which self-pollination may occur in *Eulophia* (Orchidaceae). *Journal of South African Botany*, *50*(4), 417–423.
- Withner, C. (1970). Die kleistogamen Epidendrum-arten der sektion Encyclia. Die Orchidee, 21, 4–7.
- Wong, K., & Sun, M. (1999). Reproductive biology and conservation genetics of *Goodyera procera* (Orchidaceae). *American Journal of Botany*, *86*(10), 1406–1413.

- Xiaohua, J., Dezhu, L., Zongxin, R., & Xiaoguo, X. (2012). A generalized deceptive pollination system of *Doritis pulcherrima* (Aeridinae: Orchidaceae) with non-reconfigured pollinaria. *BMC Plant Biology*, 12(1), 1–8.
- Zhang, W., & Gao, J. (2017). Multiple factors contribute to reproductive isolation between two coexisting *Habenaria* species (Orchidaceae). *PloS One*, *12*(11), e0188594.
- Zhou, X., Lin, H., Fan, X.-L., & Gao, J.-Y. (2012). Autonomous self-pollination and insect visitation in a saprophytic orchid, *Epipogium roseum* (D. Don) Lindl. *Australian Journal of Botany*, 60(2), 154–159.
- Zimmerman, J. K. (1991). Ecological correlates of labile sex expression in the orchid *Catasetum viridiflavum. Ecology*, *72*(2), 597–608.
- Zimmerman, J. K., & Aide, T. M. (1989). Patterns of fruit production in a neotropical orchid: Pollinator vs. Resource limitation. *American Journal of Botany*, *76*(1), 67–73.

# Appendix C

# APPENDIX C

## C.0.1 List of questions developed for questionnaire

**Q1:** If any species of fish live in a dark cave for several generations they eventually lose their eyesight. Describe why this may happen over time.

Q2 Fall: Why is there wood?

Q2 Spring: Wood has evolved repeatedly in many different groups of plants. Why is this?

Q3: If penguins are unable to fly, why do they have wings? Are these "wings" an adaptation?

**Q4 Fall:** Why do birds of the same species sometimes have different color coats? **Q4 Spring:** Why do male birds of the same species sometimes have different color coats?

**Q5:** If natural selection brings species closer to an optimum, then why are members of the same species not all exactly alike?

Q6: Why are there so many different species of flowering plants?

**Q7 Spring:** Some species are found in highly localized areas (endemic), while others are found across the globe (cosmopolitan). For example, *Dubautia* is a genus of flowering plants found only in the Hawaiian islands, while plants in the genus *Ranunculus* are found on every continent on earth. Why are some species endemic and some species cosmopolitan?

Q8: How do beneficial traits arise in a population of organisms?

**Q9 Fall:** Why do all mammals have four limbs? **Q9 Spring:** Why do all land mammals have four limbs?

**Q10:** Instead of having one species perfectly adapted to life on earth, there are millions of different species. Why is this the case?



C.0.2 IRB Approval #2020-0881

**Exemption Granted** 

July 30, 2020

Lucy Delaney Biological Sciences

# RE: **Protocol # 2020-0881**

"The Logic of Adaptation: How do students reason about natural selection and fit of form and function across instruction levels?"

Consistent with institutional mandates regarding COVID-19 precautions, an administrative hold has been placed on all UIC human subjects research meeting the following criteria:

- 1. The research is not designed for therapeutic benefit; and
- 2. The research involves in-person interactions with investigators or the public.

PIs who wish to begin or resume research involving activities that have been placed on temporary hold by the University due to the COVID-19 pandemic (i.e., non-therapeutic, inperson research) must complete a COVID-19 Human Subjects Research Restart Worksheet (<u>https://uic.infoready4.com/#applicationForms/1817478</u>) for an assessment of their studies prior to resuming or initiating the research.

Please refer to the Human Subjects Research Restart page on the OVCR website for additional information: <u>https://research.uic.edu/researchrestart1/</u>

The research restart is being managed by the Office of the Vice Chancellor for Research (OVCR) and the UIC Center for Clinical and Translational Sciences (CCTS). Questions about the campus research restart may be directed to research@uic.edu.

Dear Lucy Delaney:

Your Claim of Exemption was reviewed on **July 30, 2020** and it was determined that your research meets the criteria for exemption as defined in the U.S. Department of Health and Human Services Regulations for the Protection of Human Subjects [45 CFR 46.104(d)]. You may now begin your research.

Exemption Granted Date:July 30, 2020Sponsor:None

# The specific exemption category under 45 CFR 46.104(d) is: 1

You are reminded that investigators whose research involving human subjects is determined to be exempt from the federal regulations for the protection of human subjects still have responsibilities for the ethical conduct of the research under state law and UIC policy.

# Page 1 of 2

201 AOB (MC 672) 1737 West Polk Street Chicago, Illinois 60612 Phone (312) 996-1711



Please remember to:

- → Use your research protocol number (2020-0881) on any documents or correspondence with the IRB concerning your research protocol.
- → Review and comply with the <u>policies</u> of the UIC Human Subjects Protection Program (HSPP) and the guidance <u>Investigator Responsibilities</u>.

We wish you the best as you conduct your research. If you have any questions or need further help, please contact me at (312) 355-2908 or the OPRS office at (312) 996-1711. Please send any correspondence about this protocol to OPRS via <u>OPRS Live</u>.

Sincerely, Charles W. Hoehne, B.S., C.I.P. Assistant Director, IRB #7 Office for the Protection of Research Subjects

cc: Eric Stabb Joel Brown

201 AOB (MC 672)

1737 West Polk Street Chicago, Illinois 60612
# Appendix D

# APPENDIX D

## D.0.1 Font attribution

Heliotrope typeface designed by Matthew Butterick of MB Type (https://mbtype.com/).

## D.0.2 Color palette

Thesis color palette inspired by Chance the Rapper's mixtape Acid Rap.

# VITA

# LUCY E DELANEY

Professional Appointments	
Assistant Teaching Professor at University of California, Riverside Starting December Department of Evolution, Ecology, and Organismal Biology	ber 2021
Education	
Ph.D. in Ecology ନ୍ୱ Evolutionary Biology, University of Illinois at Chicago M.A. in Molecular ନ୍ୱ Cellular Biology, Hunter College of the City University of New York B.S. in Forensic Molecular Biology, Philosophy, John Jay College of the City University of New	York
Teaching Honors	
UIC Graduate Student Excellence in Teaching and Mentoring Award ( <i>honorable mention</i> )	2021

	2021
🗱 Biological Sciences Department Service Award for efforts to aid in online instruction	2021
Biological Sciences Department Graduate Teaching Award for Genetics	2020
Je Biological Sciences Department Graduate Teaching Award for Evolution	2018

## UNIVERSITY TEACHING

Instructor of Record for Foundations for the FutureFall 2021 Image: ColloquiumCourse Coordinator for Biology ColloquiumFall 2021 Image: ColloquiumCourse Builder & Trainer for the Biological Sciences DepartmentSummer 2020–Summer 2021 Image: ColloquiumGraduate Teaching Assistant for the Biological Sciences DepartmentFall 2016–Summer 2020 Image: Colloquium

## Conference Presentations and Seminars

## **Evolution Education Research**

August 2021 at University of California, Riverside | Virtual Seminar 🖄 🖓 July 2021 at Society for the Advancement of Biology Education Research | Virtual Talk 💷 🖓 June 2021 at Society for the Study of Evolution | Virtual Talk 💷 🗳 🖓 March 2021 at Midwest Ecology and Evolution Conference | 🧟 Best Graduate Talk 💷 🗳 January 2021 at SABER West | Virtual Roundtable 🗳 🖓

## Flowering Plant Breeding Systems

July 2021 at Botanical Society of America | Virtual Talk ■• 🗠 🖓 July 2018 at Botanical Society of America | Rochester, MN | Poster 皆 August 2017 at microMORPH | Arnold Arboretum of Harvard University | Talk ピ

#### PUBLICATIONS

- **Delaney, Lucy E**. (2012). Nietzsche, nerve stimulation-image connection, and ontology. *John Jay's Finest, 27*, 99–103.
- **Delaney, Lucy E** & Igić, B. (2022). The phylogenetic distribution and frequency of self-incompatibility in Fabaceae. *International Journal of Plant Sciences*. In press.

### In preparation

Delaney, Lucy E. The orchids and their breeding systems. *Prepared for Orchids magazine*.

- **Delaney, Lucy E** & Brown, J. S. University students' explanations of adaptation. I. A framework for systematic analysis.
- **Delaney, Lucy E** & Brown, J. S. University students' explanations of adaptation. II. A framework for pedagogical explanation.

## **PROFESSIONAL ACTIVITIES**

🚁 Participant in Undergraduate Mentoring Program at the Annual Evolution Conference	2021
Finvited member of UIC advisory group concerning graduate teaching support	2021
$\mathscr{F}$ Participant in the 2020 Chicago $\mathbf{Q}$ Collaborative Conference $\mathscr{O}$	2020
Reviewer for International Journal of Plant Sciences (1x), Oxford Bibliographies (1x)	2018
$\mathscr{F}$ General horticulture volunteer at Garfield Park Conservatory $\mathscr{O}$	2016

## OTHER AWARDS

🔏 Recipient of the Biological Sciences Department Travel Award	2018
🗱 Accepted to NSF-funded workshop on Bayesian Analysis of Macroevolutionary Mixtures 🔗	2017
🗱 Recipient of the Biological Sciences Department Travel Award	2017
🚁 Accepted to microMORPH Plant Anatomy Summer Course at Harvard University 🔗	2017

#### Selected Professional Experience

Forensic Molecular Biologist at NYC Office of Chief Medical Examiner	2012-2015
Health Research Intern at NYC Department of Health	2011-2012
<b>Field Manager &amp; Administrative Assistant</b> at Working Families Party New York State political action organization focused on economic justice.	2008-2011

#### Other positions

2007-2020

Like so many of our students, I held a variety of positions to support myself throughout my scholarly career: cashier, hostess, waitress, receptionist, babysitter, dog walker, bowling alley mechanic, substitute teacher, tutor, and barista. Such experiences inform my mentorship of students that face the unique challenges inherent in balancing work with the pursuit of postsecondary degrees.