

Gene discovery for comparative biology of parasitic and non-parasitic plants. A five-week molecular research immersion.

Mark A. Schoenbeck

Department of Biology

University of Nebraska at Omaha

Omaha, Nebraska 68182-0040

mschoenbeck@mail.unomaha.edu

Abstract

As part of an NSF-funded initiative, biology students, primarily those in the early stages of their degree program, were recruited to participate in ongoing research alongside principle investigators. In this example, ten students were given five weeks of intensive training in molecular research methods, with the objective of cloning genes involved in core metabolic processes in the parasitic plant *Cuscuta pentagona* (dodder) and its non-parasitic relative *Ipomoea hederacea* (ivy-leaf morning glory). During the session, students used bioinformatic tools to search gene databases, designed PCR primers, isolated genomic DNA for use in amplification and molecular cloning, mapped candidate clones, and prepared templates for sequencing. The students were provided resulting sequence data and asked to determine whether the effort had been successful. Each of these activities provided an opportunity for students to learn biological principles through their application towards a research objective.

Key words: Undergraduate research, molecular biology, bioinformatics, gene discovery, *Cuscuta pentagona*.

Introduction

This paper describes a summer research program designed in response to a solicitation for faculty involvement in undergraduate research. The STEM Talent Expansion Program (STEP), a National Science Foundation-funded initiative at the University of Nebraska at Omaha, supports faculty who provide opportunities for undergraduate participation in original research in the STEM disciplines: Science, Technology, Engineering, and Math. STEP was itself undertaken to address several aspects of undergraduate education in the STEM disciplines, including a perennial concern that undergraduates -- in this instance, undergraduates in biology and allied disciplines -- often complete their degree programs without significant research experience. Consequently, these students may lack an appreciation for the process by which facts are acquired, and the intellectual rigor, in the form of careful experimental design, integral to this process. An additional aim of STEP was to increase the rates of undergraduate recruitment into the STEM disciplines, targeting students who might be highly capable, but yet undecided about their educational goals, providing them with an opportunity to experience the daily work of a research scientist. This “second tier,” as described by Tobias (1990), represents an attractive target for STEM discipline recruitment.

The undergraduate research experience has been promoted as a means of achieving several desired outcomes, not the least of which, to promote the graduation rate

of high-quality research scientists (National Research Council, 2003). Studies of the efficacy of undergraduate research programs (reviewed by Rodrick and Dickmeyer, 2002; Seymore, et al., 2004), suggest the following benefits: an increased capacity for critical thinking; improvement in problem identification; development of technical skills; an appreciation of theory and research; and better informed decisions about graduate school.

The five-week program described in this work was similar to a conventional summer course insofar as it kept a regular daily schedule. However, the program was distinct from typical laboratory courses in that each day's activities were dictated by a single research objective, the development of which included solicited student input. Consequently, the students learned several essential techniques in the context of an integrated research program. Additionally, while some undergraduate programs offer "capstone" research courses for advanced students, this program was developed for students at the earliest stages of their undergraduate education, in the hopes that it would have a positive impact on their subsequent educational progress.

Program objectives

The pedagogical goal of the five-week research immersion was to provide students with the set of technical and intellectual skills outlined here.

Objective 1. Technical competency. Students were to master a basic set of laboratory skills. In the pursuit of the subsequently-described research objective, students would need to employ basic bioinformatic skills, accurate micropipetting, and standard

molecular techniques including nucleic acids purification, the polymerase chain reaction, gel electrophoresis, molecular cloning, and DNA sequencing.

Objective 2. Comprehension of principles underlying the methods. To compliment the development of technical skills, students were to develop an appreciation of the capacities and limitations of each. The employment of a range of technical methods provided several opportunities to present the underlying chemical and biological theory. Provisioned with this knowledge, students were able to participate in the experimental design process.

Objective 3. Comprehension of the research process and strategy. Students were asked to consider how research projects originate. Given a biological system, the students were to identify features amenable to measurement and experimentation, and that would provide information about the function of the system. Having identified a question, the students would next consider how available methods might be brought to bear in answering the question.

Recruitment of participants

Recruiting efforts were intended to target biology undergraduates at the earliest stages of their program. Invitations were extended during introductory-level botany and zoology lectures mid-way through the spring semester. As an incentive, participants could use the research to fulfill biology course electives. Non-biology undergraduates were also encouraged to participate, and a similar invitation was extended through the

Principles of Biology course, which largely serves this audience. A majority of participants identified themselves as biology, biotechnology, and environmental studies majors, though students from English, math, and computer science programs also enrolled. Each summer's research group comprised at least ten students working in pairs. Student surveys conducted at the beginning of the session indicated that, with the rare exception, none of these students had participated in an organized research effort prior to this program.

The parasite dodder, and its non-parasitic relative, ivy-leaf morning glory.

While the research program had the previously-stated pedagogical objective of training the participants to function as capable researchers, the daily laboratory activities were focused on achieving an explicit research objective. The research focus was the comparative physiology of a parasitic organism with a non-parasitic relative, in this instance two flowering plant species. After comparing the growth and developmental habits of the parasite and its relative, the students chose for further investigation genes that likely contribute to physiological functions that might have changed in the process of, or subsequent to, the adoption of the parasitic habit. The students would then attempt to isolate gene fragments from the corresponding genes of both species, to serve as tools for subsequent molecular investigations.

Cuscuta pentagona, shown in Figure 1a, is a rootless, leafless, epiphytic parasite representative of the genus commonly called the dodders. Emerging from seed, the dodder seedling uses environmental cues such as phytochemical emissions (Runyon et al., 2006) or light quality (Furuhashi et al., 1995; Haidar et al., 1997) to seek out potential

host plants and establish parasitic attachment. Through a combination of nutation and directional growth, the seedling twines around any object it encounters. Upon contacting most plant tissues, and also many non-plant materials, dodder uses a secretion to adhere to the target, subsequently breaching the host surface with penetrating haustoria that grow into an association with the host phloem (Vaughn, 2002; Vaughn, 2003). It is from the host phloem then, that the parasite appears to draw most of its fixed carbon and presumably all of its nitrogen and mineral resources (Jeschke et al., 1994). Current analyses indicate that the clade Cuscutaceae has emerged from the Convolvulaceae (Neyland, 2001; Stephanovic and Olmstead, 2004), and morning glory species have previously been used as points of comparison in characterizing the biology of the parasite (MacPherson, 1921.)

Since the completion of the Arabidopsis genome project (Arabidopsis Genome Initiative, 2000) most of the genes encoding enzymes of plant core metabolism have been identified. For the purpose of this investigation, “core metabolism” comprises the major reactions the Calvin cycle, starch synthesis and breakdown, glycolysis, sucrose synthesis, and organic acid metabolism as it relates to the primary assimilation of ammonium into glutamine, glutamate, aspartate, and asparagine, as illustrated in Figure 2. Given the developmental and trophic differences between the parasite and non-parasite, students were asked to postulate how expression of the genes underlying these processes might differ. As an example, many non-parasitic plants transport environmentally-acquired mineral nitrogen, in the form of nitrate, from the roots to the leaves via xylem, at which point the nitrate is reduced to ammonia and assimilated into the amino acid pool. Lacking roots and leaves, and being attached primarily to its host’s phloem, it is unclear

under what circumstances *Cuscuta* might employ nitrate reductase, the first step in nitrate reduction. While nitrate reductase activity is not detected stems of parasitizing plants, it may be detected under artificial conditions following nitrate application (Chauhan and Srivastava, 1980; Schoenbeck, unpublished observations). Furthermore, the activity of ammonia-assimilating enzymes, such as glutamate dehydrogenase, may also be detected under some circumstances (Srivastava and Dwivedi, 2003). The genes encoding the enzymes of photosynthesis and carbon metabolism could be considered in a similar light. The identification of these genes, and those encoding other activities indicated in Figure 2, in both the parasite and the non-parasitic relative, would provide a starting point for the comparison of how these functions are employed.

Overview of activities.

The progress of research activities is outlined in Figure 3. The first two weeks were dedicated to establishing technical competency and to the development of a research plan for the recovery of the target genes. The remaining three weeks were used for implementation of the strategy and evaluation of results.

Development of technical skills. As nearly all aspects of molecular cloning require the precise handling of small reagent volumes, a significant effort was dedicated to training the students in accurate micropipetting. Students were introduced to the range of micropipetter that they would be using, and instructed in how to read and adjust them. Emphasis was placed on “chemical hygiene” pointing out the ease with which cross-contamination of reagents could occur should they fail to use clean tips, or accidentally

draw materials up into the pipetter barrel. Students practiced their skills in liquid handling by pipetting volumes of liquids with different viscosities and with low surface tension (isopropanol and chloroform) to observe the limitations of a negative-displacement pipetter, and practices employed to overcome these difficulties, such as pre-saturating the tip volume with solvent vapor prior to pipetting, in order to prevent the liquid from running out. An especially helpful exercise had the students generated a row of 5 μL and 1 μL droplets on a strip of Parafilm using a dye solution, comparing their volumes to an example placed by the instructor or lab assistant. Using this approach, the students quickly came to recognize the volumes they were to be working with regularly.

Pipetting error was the most frequently identified cause for control experiments to fail. Initially students, working as teams, were issued freezer boxes with volumes of molecular reagent that should have been sufficient for several days' work. However it was soon recognized that students' unfamiliarity with typical working volumes could result in their using several-fold excess of a reagent. Improved performance was achieved in subsequent sessions by issuing measured amounts of reagents, sufficient for the experiment at hand, and explaining to the students that they had a limited margin for error in manipulating the correct volumes. Pipetting skills increased markedly under these strictures, as students monitored their pipetting much more closely and identified their errors sooner. Upon mastering pipetting skills, students practiced casting and loading agarose gels for DNA electrophoresis. These exercises were repeated until the most frequent mistakes and poor technique were identified and eliminated.

A third technical skill developed during the first two weeks of the session and central to the conduct of the research was the polymerase chain reaction (PCR). The

principles of PCR were presented from the perspective of DNA replication, noting all of elements needed for synthesis of a new DNA strand: DNA polymerase, nucleotides, template, and a free DNA 3' end as a starting point. Subsequently, the features that make PCR useful for experimental purposes were introduced -- that the location of the 3' "priming" site could be determined by designing artificial oligonucleotides to anneal to selected points on the template, and that repeated cycling in a reaction with oppositely-oriented primers resulted in exponential synthesis of the region between them.

As previously noted, students were initially issued a standard kit of PCR reagents, including a stock of dNTPs, MgCl₂, 10X concentrated PCR buffer, and Taq DNA polymerase. Using a stock template of plasmid and standard primers, the success rate of initial attempts was typically less than half of the class. With second and third attempts, students generally became confident in their ability to assemble a successful reaction. However, the amount of pipetting required was clearly the factor that led to the high variability in results. In subsequent sessions, students were issued a stock of PCR "ready-mix", comprising buffer, dNTPs, and MgCl₂ at working concentrations, and requiring only the addition of primers, template, and Taq polymerase. After substituting this "ready-mix", the rate of student success at PCR on the first attempt increased to greater than half of the class, and several students were able to generate consistently successful amplifications.

At stages following the second week, students also practiced assembling DNA ligation reactions, transforming recombinant plasmid vectors into *Escherichia coli*, and recovering plasmid DNA from bacterial cell cultures, skills that would be required in the later stages of the project.

Developing a strategy

At the same time that students were developing their molecular research skills, they were also engaged in developing a research strategy for examining the physiological differences between the parasite dodder and its non-parasitic relative through the identification of genes that may contribute to these processes. The students were guided in this task through a series of discussions.

Introduction to the process of research. To introduce students to the process of research, they were asked to offer their best understanding of where and why research is performed, and by whom. While most students had a general appreciation that research occurs at universities and similar institutions, they were largely unaware of the typical structure of a laboratory regarding the roles of the principle investigator, graduate and undergraduate researchers, post-doctoral scholars, and technicians. The description of the principle investigator's role in providing the focus and direction for research led to the question "what constitutes a good research objective, and how does one begin to pursue it?"

The students were given for consideration the contrasting habits of dodder and morning glory, and asked to propose what biological insight might be gained by studying the differences between these organisms. Through many observations and suggestions, a theme emerged, addressing the question of how core physiological processes might change in the course of, or as a consequence of, the adoption of the parasitic habit. The students were subsequently introduced to the central dogma of molecular biology,

relating cellular functions, mediated by enzymes and other proteins, to distinct genes. Consequently, the students were able to propose that the identification and characterization of genes for core metabolic processes could serve as a promising starting point for understanding how these two very different plants function. A list of enzyme activities contributing to the major metabolic pathways, as illustrated in Figure 2, thus provided a list of cognate genes for subsequent examination.

Establishing a research plan. A set of research skills that were clearly not well-developed among the student researchers was that needed for the identification, gathering, and employment of current knowledge in developing a research strategy. Simply stated, even given the objective, there was little sense of how to proceed. It was thus impressed upon them that it is important to establish the “state of the science” concerning the research objective, as learning how other researchers have addressed a problem will provide ideas for how to advance a new project. In this example, the question at hand was whether the target genes had been identified previously, whether in the target plant species, or from other closely- or distantly-related plant species. The pursuit of the answer to this question was the focus for an introduction to commonly-used bioinformatic tools.

The National Center for Biotechnology Information site (<http://www.ncbi.nlm.nih.gov/>) hosts an array of tools that may be employed to help students develop competency in searching for relevant information. Information was gathered using a combination of search strategies. Initially, students used the PubMed search function to identify helpful publications and their authors, submitting

combinations of key words such as the enzyme activity and the organism. Once it had been established that a gene for a target enzyme had been identified, the author name could be used to search the nucleotide database and find the cognate sequence. An alternative approach was to use the Taxonomy database to determine how many nucleotide sequences were available from the target species, genus, or family, noting whether any of the known, or putative, genes were among those contributing to core carbon or nitrogen metabolism.

Using a known nucleotide sequence as a starting point, the students used the Basic Local Alignment Search Tool (BLAST, Altschul et al., 1990) to identify similar gene sequences from other sources. When they had retrieved gene sequences from a range of species, including both species closely related to Convolvulaceae, and species more distantly related. Sequence alignments were generated using the Clustal X DNA alignment program (Thompson et al., 1994). Figure 4 illustrates the sequence alignment between starch synthase genes from members of the clade Solanales (to which Convolvulaceae and Solanaceae both belong), comprising sweet potato, tomato, and tobacco, and more distantly-related Arabidopsis. Because many of the students had only limited exposure to advanced concepts of molecular genetics, this gene alignment exercise served as an effective method for demonstrating several characteristics of the genetic code:

Regions of high and low sequence identity – Students noted that not all regions of aligned gene sequences share the same degree of sequence identity. Speculating that this reflected a difference in selection pressure among different locations within the gene,

students could thus choose the most highly-conserved regions of the gene as potential sequences for DNA oligonucleotides for use in PCR.

Gene structure is conserved between organisms – When more than one genomic sequence was available for a given enzyme-coding gene, students noted that the position of introns and exons tended to be conserved. This was important in that, as many available sequences were from cDNA clones lacking introns, students could predict the location of introns and avoid the misplacement of primers across spliced exon-exon junctions. As part of the experimental design, students were encouraged to design their primer pairs in such a way as to “capture” at least one intron from the target species’ gene. In this way they could test a prediction about the conservation of gene structure.

The nucleotide triplet nature of the genetic code, and variability at the third nucleotide position in the codon – DNA alignments repeatedly showed non-conserved nucleotides occurring at intervals of three positions, or multiples thereof, making it a simple matter to surmise the reading frame of the gene. Single nucleotide deletions at one position, resulting in a frame-shift relative to aligned sequences, were often seen to be compensated by nearby deletions or insertions that restored the reading frame.

Using the parsimony principle in comparing molecular data – Students found that there were frequently regions of conserved sequence identity in all but one of the aligned sequences (see Figure 4). While these regions of generally conserved sequence were compelling sites for primer design, the question was posed, what was to be done with the

nucleotide at the non-conserved site? Options such as the use of degenerate primers were proposed, but the students were also asked to consider, given the relationship between the species compared in the sequence alignment, whether it might be possible to make an educated guess as to the most likely nucleotide. For example, when several of the sequences from plant species more closely related to the target species share a sequence that is not shared with more distantly related species, the parsimony principle would hold that the shared sequence is a shared trait among the near relatives, representing the “best educated guess” for the matching sequence in the target genome. Employing the same line of reasoning, if a single member of the near-relative clade possesses a distinct sequence, while all other near and distant relatives share a common sequence, the non-consensus individual would be considered “derived”, while the others would be said to have the “ancestral” state.

By end of the second week of the session, the students had submitted their suggestions for oligonucleotide primer combinations. The suggested primers were considered in a group discussion, noting the degree of sequence identity at each site and the predicted size of the amplified product. The most promising combinations were submitted for synthesis. Typically, multiple forward and reverse primer combinations were generated for each gene that could be used combinatorial experiments.

Preparation of research materials (week three)

During the third week, student researchers prepared the necessary genomic DNA for use as the template for experimental PCR using one of several suitable

phenol:chloroform-based DNA extraction methods (Nishiguchi et al., 2002). Before proceeding, the students were led through the protocol in detail, explaining the chemistry of nucleic acid purification and the purpose of each stage, from cell disruption to the final spooling of precipitated DNA. Students first practiced DNA isolation from leaf tissue of *I. hederacea*. The quality and quantity of the prepared DNA were judged by inspection of a sample following gel electrophoresis. Each student prepared a minimum of two DNA samples, and each team was required to have at least one good DNA preparation from each *C. pentagona* and *I. hederacea*.

Experimental amplification, cloning, and screening (week four)

With suitable templates in hand by the end of the third week, students were ready to attempt the amplification of target gene fragments from each of the two species. As previously noted, when sequence identity between aligned gene sequences permitted, multiple forward and reverse primers were designed for each target gene, so that multiple combinations of primers were available. For any given experimental PCR, the student teams were allowed to choose the primer combination and the template that they would pursue. Students were required to prepare all needed controls (forward primer alone with template, reverse primer alone with template, forward and reverse primers together but without template) to qualify any resulting amplification products. Results of a typical student gel are presented in Figure 5a. Success rates varied between groups, but in every instance, there were enough successful amplifications – yielding amplification products in the experimental, as opposed to the control reactions – to provide material for all groups to participate in subsequent steps.

Amplified products were ligated into the pGEM PCR cloning vector (Promega) and transformed into *E. coli* strain DH5 α plated onto solid LB medium containing ampicillin for selection. As a fraction of the ligation reaction regularly generated closed plasmids lacking inserts, this exercise provided an opportunity to distinguish between *experimental selection* for bacteria carrying the ampicillin resistance gene on the vector, and *experimental screening*, employing blue versus white colony color developed in the presence of X-gal, distinguishing plasmids without inserts from plasmids with inserts, respectively. White bacterial colonies were screened for insert size; students picked bacteria directly from the plate into PCR mix with primers flanking the plasmid cloning site. In this way the students could rapidly determine which colonies were carrying fragments of a size comparable to candidate amplification products from the original experimental amplification.

Molecular characterization of cloned fragments (weeks four and five)

From among the bacterial colonies with plasmids containing an insert, student researchers were to choose the best candidates, based primarily on size predictions for the corresponding fragment in known sequences, for subsequent analysis.

Clone mapping- This first step in the characterization of individual clones was the mapping of the inserted fragment's orientation with respect to the vector. Students used the forward primer from the experimental amplification in combination with either of the two primers flanking the cloning site, using these as reference points. Representative results of a student's mapping exercise are shown in Figure 5b and 5c.

Sequencing and analysis- From the results of the screening and mapping exercises, student could choose cloned fragments that represented the most promising candidates, those that matched predictions for the size of the targeted gene region. Each group generated template DNA by amplification of the clone from T7 and Sp6 priming sites flanking the cloning site. PCR-amplified template was cleaned using a glass-powder protocol, and, upon confirming sufficient template recovery, students assembled sequencing reactions (Sequi-Therm sequencing reagent, Epicentre) using fluorescent tag-labeled T7 primer. To the extent that time permitted, students assisted in sequencing gel preparation and sample loading on a Licor 4300 DNA analyzer. Alternatively, prepared samples were submitted to an external sequencing facility. DNA sequence data resulting from the sequencing reactions was returned to the students directly, and it became their responsibility to determine, by comparing their sequence data with database sequences, whether they had been successful in their pursuit.

A majority of the student-generated clones were determined to have sequences that did not match the target gene. However, over the course of three summers' effort, students were successful in recovering cloned fragments with sequence similarity to plant aspartate aminotransferase, glyceraldehyde 3-phosphate dehydrogenase, and starch synthase genes, from one or both of the target organisms. These clones are being used as molecular probes in the comparison of cognate gene expression levels between the target species. Other genes that were pursued without success included malate dehydrogenase, phosphoenolpyruvate carboxylase, and nitrate reductase.

The wrap-up

At the conclusion of the research session, the group completed surveys addressing the quality of the instructor and the research activities. Uniformly, the students indicated that they perceived the research experience to have been positive, that increased their comprehension of some of the more complicated concepts of genetics and molecular biology, and for those majoring in biology, that they expected to benefit from the research in their subsequent coursework. Communications with students one or more years after their participation has indicated that these students have especially benefited in their increased capacity to operate with confidence in a laboratory setting.

Notes on the pace of daily work. With the exception of those exercises requiring a computer lab, all of the activities reported here were performed in a standard science teaching lab, during a three hour block, four days per week, with one additional day reserved for necessary preparation. As the daily worked often entailed a waiting period, as in the case of running gel electrophoresis or PCR, the discussion of research strategy were interjected at opportune times. The amount of work here described, including the time spent training students, consumed the entire five-week schedule.

Acknowledgements

Portions of this project were funded from the NSF STEP grant, NSF-0336462 and NIH Grant Number P20 RR16469 from the INBRE Program of the National Center for Research Resources.

Literature cited

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410.

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815.

Chauhan, J.S. and Srivastava, H.S. 1980. Some aspects of regulation of nitrate reductase activity in the seedlings of an angiosperm parasite: *Cuscuta reflexa* Roxb. *Plant Biochemical Journal* 7, 25-36.

Furuhashi, K., Kanno, M., and Morita, T. 1995. Photocontrol of parasitism in a parasitic flowering plant, *Cuscuta japonica* Choisy, cultured in vitro. *Plant and Cell Physiology* 36, 533-536.

Haidar, M.A., Orr, G.L., and Westra, P. 1997. Effects of light and mechanical stimulation on coiling and prehaustorial formation in *Cuscuta* spp. *Weed Research* 37, 219-228.

Jeschke, W.D., Rath, N. and Bäuml, P. 1994. Modeling the flow and partitioning of carbon and nitrogen in the holoparasite *Cuscuta reflexa* Roxb. and its host *Lupinus albus*

L. II. Flows between host and parasite and within the parasitized host. *Journal of Experimental Botany* 45, 801-812.

MacPherson, G.E. 1921. Comparison of development in dodder and morning glory. *Botanical Gazette* 71, 392-398.

National Research Council. 2003. *Bio 2010. Transforming undergraduate education for future research biologists.* National Academies Press, Washington, D.C.

Neyland, R. 2001. A phylogeny inferred from large ribosomal subunit (26S) rDNA sequences suggests that *Cuscuta* is a derived member of Convolvulaceae. *Brittonia* 53, 108–115.

Nishiguchi, M.K., Doukakis, P., Egan, M., Kizirian, D., Phillips, A., Prendini, L., Rosenbaum, H.C., Torres, E., Wyner, Y., DeSalle, R., and Giribet, G. 2002. DNA Isolation Procedures. *In Methods and Tools in Biosciences and Medicine*, pp. 250-287. R. DeSalle et al., eds. Birkhäuser Verlag Basel/Switzerland.

Rodrick, R., and Dickmeyer, L. 2002. Providing undergraduate research opportunities for communication students: A curricular approach. *Communication Education* 51, 40-50.

Runyon, J.B., Mescher, M.C., and DeMoraes, M.C. 2006. Volatile chemical cues guide host location and host selection by parasitic plants. *Science* 313, 1963-1967.

Seymore, E., Hunter, A.B., Laursen, S.L. and Deantoni, T. 2004. Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education* 88, 493-534.

Srivastava, S. and Dwivedi, U.N. 2003. Modulation of key nitrogen assimilating enzymes by NAA and in vitro culture in *Cuscuta reflexa*. *Plant Physiology and Biochemistry* 41, 65-71.

Stephanovic, S. and Olmstead, R.G. 2004. Testing the phylogenetic position of a parasitic plant (*Cuscuta*, Convolvulaceae, Asteridae): Bayesian inference and the parametric bootstrap on data drawn from three genomes. *Systematic Biology* 53, 384-399.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-80.

Tobias, S. 1990. *They're Not Dumb, They're Different: Stalking the Second Tier*. Tucson: Research Corporation.

Vaughn, K.C. 2002. Attachment of the parasitic weed dodder to the host. *Protoplasma* 219, 227-237.

Vaughn, K.C. 2003. Dodder hyphae invade the host: a structural and immunocytochemical characterization. *Protoplasma* 220, 189-200.

Figure legends

Figure 1. The parasitic angiosperm *Cuscuta pentagona* on its host (left) and its non-parasitic relative *Ipomoea hederacea*. Recent molecular investigations place the genus *Cuscuta* in the family Convolvulaceae with *Ipomoea*.

Figure 2. A generalized scheme of carbon and nitrogen core metabolic processes in C3 plants. Shaded boxes indicate physiological processes that are necessarily reduced, or possibly absent, in species of *Cuscuta*. Of physiological interest is the determination of which shaded processes are indeed present and to what degree, or to determine how related processes (unshaded) are affected by the parasite's altered metabolism. Enzymes contributing to core metabolism are indicated by gray boxes: AAT, aspartate aminotransferase; AS, asparagine synthetase; ALD, aldolase; AMY, amylase; GDH, glutamate dehydrogenase; GOGAT, glutamate synthase; GS, glutamine synthetase; MDH, malate dehydrogenase; NR, nitrate reductase; NiR, nitrite reductase; PEPC, phosphoenolpyruvate carboxylase; RUBISCO, ribulose 1,5-bisphosphate carboxylase/oxygenase.

Figure 3. A flow chart describing the strategy for identifying genes involved in core metabolic processes in dodder and ivy-leaf morning glory. The gene discovery strategy employed a range of skills, comprising the use of bioinformatic tools as well as standard molecular research methods.

Figure 4. Partial sequence alignment of plant starch synthase sequences used for PCR primer design. Starch synthase cDNA sequences from sweet potato (*Ipomoea batatas*, Genbank accession gi:15637078), tomato (*Lycopersicon esculentum*, gi:47104845), tobacco (*Nicotiana tabacum*, gi:6116747) and *Arabidopsis thaliana* (athalc, gi:23506180) were saved into a plain-text file and formatted for use with the Clustal X sequence alignment program. The genomic sequence corresponding to the *Arabidopsis* cDNA, locus At1g32900 (athalg), was also included to provide a reference for probable intron locations within the gene. Dashes correspond to the location of intronic sequence as predicted by the *A. thaliana* genomic sequence. The dotted box indicates a region of high sequence conservation, a good candidate for primer design, while the dotted line indicates sequence data omitted for brevity in this figure. Site A indicates a substitution where it is possible to choose a nucleotide based on its common occurrence in plants across the order Solanales, though it does not share identity with *Arabidopsis*. Site B indicates a substitution where the ancestral state is evident, and substitution has occurred more recently in a particular solanaceous lineage. Site C indicates a site where there it is not possible to determine the ancestral state. The forward primer designed by the group is underlined.

Figure 5. Student amplification of putative starch synthase gene fragment from *Cuscuta pentagona*, PCR mapping of a cloned fragment to determine orientation, and schematic map of the cloned fragment. A. PCR amplification experiment using *C. pentagona* genomic DNA as a template and primers described in Figure 4. Primer combinations are indicated at the bottom of the gel: F=SSFwd, R=SSRev, -C=negative control in from which template DNA has been omitted. A single amplification product appears within the predicted size range in lane 3. B. Single primers and primer combinations were used to demonstrate the orientation of a cloned fragment (not the same fragment as in panel A). Primer combinations are the same as in panel A, with the addition of the T7 primer. C. The amplification pattern from B demonstrates that the T7 and SSFwd primers are in opposition to each other, indicating that the cloned fragment is in the “reverse” orientation.