Active Assessment: Assessing Scientific Inquiry
Biology is evolving rapidly, with more and more discoveries arising from interaction with other disciplines such as chemistry, mathematics, and computer science. Undergraduate and Graduate biology education is having a hard time keeping up. To address this challenge, this bold and innovative series will assist science education programs at research universities, four-year colleges and community colleges across the country and by enriching science teaching and mentoring of both students and faculty in academia and for industry representatives. The series aims to promote the progress of scientific research and education by providing guidelines for improving academic and career building skills for a broad audience of students, teachers, mentors, researchers, industry, and more.

**Volume 1**  Education Outreach and Public Engagement  
by Erin L. Dolan

**Volume 2**  Active Assessment: Assessing Scientific Inquiry  
by David I. Hanauer, Graham F. Hatfull, Deborah Jacobs-Sera
Active Assessment: Assessing Scientific Inquiry
We dedicate this book to the community of scientist-educators and to all – from both sides of the divide – who aspire to join them.
Acknowledgments

We would like to extend thanks to our many colleagues at the University of Pittsburgh that have been integrally involved in the study of bacteriophage genomics. Dr. Roger Hendrix spearheaded the phage comparative genomics effort in the early 1990s, and it was from this foundation that the PHIRE program emerged. Roger has been a truly wonderful and inspiring colleague and collaborator through all aspects of the genome investigations and the PHIRE program. We also thank Dr. Jeffrey Lawrence who has been centrally involved in the phage genome projects since joining the University of Pittsburgh, and we are grateful for his frequently profound insights into microbial evolution as well as for the enormous investment of time and effort into the development of suitable bioinformatic tools and programs. Dr. Craig Peebles has aided in phage genome analysis in numerous ways especially in the identification of small tRNA genes and other features; he has also supervised many students in the PHIRE program.

We extend special thanks to Dr. Marisa Pedulla who was the PHIRE program coordinator when the program was initiated in 2002. As a dedicated scientist/educator, Marisa established a strong foundation for the PHIRE program and established many of the structures that ensured its continuing success including introducing David Hanauer to the PHIRE program team. The PHIRE program could not have been successful without Marisa’s dedication to placing students at the heart of our phage discovery and genomic effort.

Dr. Bill Jacobs introduced GFH to the mycobacteriophage world in 1988 and has been a fantastic collaborator in every aspect of research and the PHIRE program. Bill has had undergraduate and high school phage hunters in his laboratory for many years and has been an inspiration for all of us working with mycobacteriophages.

We are fortunate to have a small genome center at the University of Pittsburgh that is dedicated to phage genomics, and over the years has had a wonderful team of research associates that have skillfully helped to sequence phage genomes and have tirelessly mentored students in genome sequence and analysis. We also thank Dr. Steve Cresawn who participated in all components of the PHIRE program, as well as designing and composing the Phamerator program for analyzing gene phamilies.

The PHIRE program and the work on this book were supported by a Professorship from the Howard Hughes Medical Institute. We are enormously grateful to the
support that HHMI has provided, both financially and programmatically. Dr. Peter Bruns was helpful and supportive in all aspects of administration and execution of the PHIRE program and we are grateful for his leadership at HHMI. We also thank Maria Kozalka at HHMI for outstanding assistance, especially at the outset of the Professors programs and when we were establishing the PHIRE program. The HHMI Professors program provided a critical opportunity to meet and learn from many truly inspiring scientist/educators. There is not one of the professors that did not provide some special insight and perspective, but we would like to especially thank Drs. Jo Handeslman, Utpal Banerjee, and Scott Strobel for their insights and inspirations.

Our graduate students, postdoctoral researchers, research associates, and fellow faculty have joyfully participated in the grand PHIRE experiment. They have assisted in ways too numerous to mention, have helped mentor many students, and have shown remarkable tolerance for all the perturbations that can arise from having relatively inexperienced researchers in the laboratory milieu.
Contents

1 Active Assessment ................................................................. 1
  1.1 Scientists as Educators ...................................................... 1
  1.2 The Context and Aims of This Book ........................................ 1
  1.3 Relevant Historical Developments in Science Education ................. 2
  1.4 Active Assessment Defined .................................................. 3
  1.5 The Underpinning Principles of Active Assessment .......................... 4
  1.6 On the Importance of Active Engagement ..................................... 6
  1.7 The Design of This Book ...................................................... 8

2 Conceptualizing Scientific Inquiry ............................................. 11
  2.1 Introduction ........................................................................... 11
  2.2 The Diversity of Scientific Inquiry ............................................ 11
  2.3 The Characteristics of Authentic Scientific Inquiry .......................... 15
  2.4 An Analytical Framework for the Definition of Scientific Inquiry ...... 17
  2.5 Chapter Summary ................................................................. 20

3 An Introduction to Assessment .................................................. 23
  3.1 Assessment and Science ......................................................... 23
  3.2 Assessment, Evaluation, and Testing .......................................... 23
  3.3 Summative, Formative, and Diagnostic Assessment ......................... 24
  3.4 Distinctions in Elicitation Methods ............................................. 25
  3.5 Developing an Assessment Program ............................................ 26
  3.6 Chapter Summary ................................................................. 29

4 Assessing Scientific Inquiry .................................................... 31
  4.1 Introduction ........................................................................... 31
  4.2 Approaches to the Assessment of Scientific Inquiry .......................... 31
  4.3 Authentic Assessment ........................................................... 37
  4.4 The Characteristics of Authentic Scientific Inquiry Assessment (ASIA) .... 39
  4.5 Active Assessment Development .............................................. 41
  4.6 Chapter Summary ................................................................. 42
5 An Analytical Framework for the Development of Scientific Inquiry Assessment ..................................................... 45
5.1 Introduction ................................................................ 45
5.2 Stages of Active Assessment Development ......................... 45
   5.2.1 Stage 1: Empirical Description of Scientific Inquiry .... 46
   5.2.2 Stage 2 – Definition of Aims ................................. 49
   5.2.3 Stage 3: Tool Development ................................ 50
   5.2.4 Rubric Development ........................................... 51
   5.2.5 Assessment Piloting ............................................. 54
5.3 Chapter Summary ..................................................... 54

6 The PHIRE Program ...................................................... 55
6.1 Introduction ................................................................ 55
6.2 What Are Phages? .................................................... 56
6.3 Bacteriophages Are Ubiquitous and Host Specific ................. 58
6.4 Bacteriophages: The Dark Matter of the Biological Universe .. 59
6.5 Goals of the PHIRE Program ....................................... 60
6.6 The Power of Discovery .............................................. 60
6.7 PHIRE: A Ten-Step Program ......................................... 61
   6.7.1 Step 1: Phage Isolation ......................................... 61
   6.7.2 Step 2: Phage Purification ..................................... 62
   6.7.3 Step 3: Phage Amplification .................................. 62
   6.7.4 Step 4: Electron Microscopy .................................. 63
   6.7.5 Step 5: Nucleic Acid Extraction and Restriction Analysis . 64
   6.7.6 Step 6: DNA Sequencing ..................................... 66
   6.7.7 Step 7: Genome Annotation ................................... 67
   6.7.8 Step 8: Comparison of the DNA Sequence to Known 
               Phage Genomes .................................................. 67
   6.7.9 Step 9: Comparative Genomic Analysis ...................... 69
   6.7.10 Step 10: Publication ............................................ 72
6.8 Is That All? .............................................................. 72
6.9 What’s in a Name? ..................................................... 73
6.10 Who’s Qualified for the PHIRE Program? ......................... 74
6.11 Parallel Projects: Pros and Cons .................................. 74
6.12 Mentees and Mentors: Opportunities and Responsibilities .... 75
6.13 Scheduling and Flexibility .......................................... 76
6.14 Multiple Milestones of Success .................................... 76
6.15 Not All Failure Is Bad! .............................................. 77
6.16 Transitions: From Concrete Beginnings to Abstract Representation . 77
6.17 Seven Attributes of the PHIRE Program .............................. 78
   6.17.1 Attribute 1: Technical Simplicity with Transition 
           to Complexity ...................................................... 78
   6.17.2 Attribute 2: Conceptual Simplicity with Transition 
           to Complexity ...................................................... 79
Contents

6.17.3 Attribute 3: Flexible Scheduling Compatibility ........... 79
6.17.4 Attribute 4: Multiple Achievement Milestones ............ 79
6.17.5 Attribute 5: Parallel Projects and Mentoring ................. 79
6.17.6 Attribute 6: Authentic Publishable Research ................. 79
6.17.7 Attribute 7: Project Ownership ................................ 79

7 The PHIRE Program Assessment Strategy .......................... 81
  7.1 Introduction ................................................. 81
  7.2 Aims of Undergraduate Phage Hunting ............................ 81
  7.3 Detailed Analyses of Decision-Making Reasoning and Calculation in the Bacteriophage Isolation and Identification Process ........ 83
    7.3.1 Phage Isolation – Identification Through Infection .......... 84
    7.3.2 Phage Purification Through Selective Infection .............. 85
    7.3.3 Phage Amplification – Measuring Bacteriophage Concentration to Determine and Produce Maximum Bacteriophage Yield ........................................ 86
    7.3.4 Electrophoresis ........................................... 87
    7.3.5 Gene Determination (Annotation) ............................ 88
  7.4 An Educational Assessment Strategy for the Scientific Inquiry Process of Phage Hunting ........................................ 94
  7.5 Weighting and Timing ......................................... 97

8 PHIRE Assessment Tools .............................................. 99
  8.1 Introduction .................................................. 99
  8.2 The Substantive Knowledge Test .................................. 99
  8.3 The Physical Checklist .......................................... 100
  8.4 The Visual Literacy Test ........................................ 101
  8.5 The Notebook Assessment Tool ................................... 103
    8.5.1 Program Beginnings and Notebook Introduction ............ 104
    8.5.2 On-Going Laboratory Notebook Usage and the Provision of Feedback ........................................ 106
    8.5.3 Final Assessment of Laboratory Notebooks (End of Program) ........................................ 107
  8.6 The Knowledge Presentation Performance Test .................... 110
    8.6.1 Conference Poster Development ............................. 111
    8.6.2 Conference Poster Presentation ............................. 112

9 Reflections from an Active Scientist-Educator ...................... 115
  9.1 Introduction .................................................. 115
  9.2 Two for the Price of One: Integrating the Missions ............. 116
  9.3 Some Practical Concerns ....................................... 117
    9.3.1 Space: The Final Frontier .................................. 117
    9.3.2 Money and Time .......................................... 118
Chapter 1
Active Assessment

1.1 Scientists as Educators

Doing science is an exciting, fulfilling activity that contributes to the collected cultural knowledge of humanity. By doing science new knowledge is created and our understanding of the world around us is increased. Unfortunately, this sense of mission and personal feelings of excitement and fulfillment so characteristic of the active scientist are lost in most science education settings. This book, as with other publications from the authors of this book, evolved from the simple premise that active scientists should be involved in and develop serious educational programs designed around real scientific research questions. Science education is too serious a purpose that scientists can leave this to others to fulfill. The authors of this book hold a deep belief that the way to advance science education is through the development of in-laboratory science education programs that bring students into close contact with the experience and realities of authentic scientific inquiry. This book is designed to help active scientists to create authentic assessments that contribute to the educational process and provide meaningful data that can be used by the scientist-educator and student-scientist to enhance the educational process within the laboratory and thus enhance science. This book was written with the understanding that scientists may not feel comfortable with educational concepts and terminology and that this is one of the barriers to the creation of more laboratory-based, authentic science education programs. It is our hope that this book will provide a clear introduction to the approach to the assessment of scientific inquiry that we have developed and used in our own laboratory educational program and that this knowledge will ultimately lead to the creation and understanding of assessment in new laboratory-based scientific inquiry educational programs.

1.2 The Context and Aims of This Book

This book deals with a very specific educational context – the in-laboratory, scientific inquiry, educational program. The issue that this book addresses is the way
to assess knowledge development and outcomes within this setting. This book has three main aims:

1. To provide scientist-educators working with in-laboratory scientific inquiry educational programs an approach to the development of a meaningful assessment program
2. To provide scientist-educators with a comprehensive understanding of issues of educational assessment and an overview of the work that has been done by science educators concerning the assessment of scientific inquiry
3. To provide scientist-educators with a case study and specific examples of one program that utilized the approach developed in this book to the assessment of scientific inquiry.

1.3 Relevant Historical Developments in Science Education

This book finds its source in three interrelated developments in science education. The first deals with the importance of enhancing students’ understanding of the procedural knowledge of scientific activity. The most widely recognized statement of this type was the National Research Council’s publication of their *National Science Education Standards* promoting scientific inquiry as a core element of scientific education (NRC, 1996). Over the years since this early publication, the same message of the centrality of scientific inquiry as an educational tool has been repeated in a variety of publications and national reports. Most importantly for undergraduate science education, in another NRC report prepared by the Committee for Undergraduate Biology Education to Prepare Research Scientists for the 21st Century and entitled *BIO2010: Transforming Undergraduate Education for Future Research Biologists*, an emphasis is placed on providing students with the experience, understanding, and skills required to conduct interdisciplinary scientific inquiries within the coming century.

The second educational development consists of the movement of science education out of the classroom and into the laboratory. This direction can be seen as the most direct implementation of the concept that science education should focus on scientific inquiry. As stated by Handelsman et al. (2004), “Scientists of all disciplines have developed inquiry-based labs that require students to develop hypotheses, design and conduct experiments, collect and interpret data, and write about results” (p. 521). Extending this argument Hatfull et al. (2006) and Hanauer et al. (2006) promote (and exemplify) the transformation of the professional research laboratory into an educational arena that brings students from a diversity of disciplines, ages, and interests into a laboratory in order to conduct authentic scientific inquiries. This requires the definition of in-lab research projects that are appropriate for a variety of incoming students with different knowledge bases but still culminate in authentic scientific discoveries (Hatfull et al., 2006; Hanauer et al., 2006).
The third educational development deals with the role of the science instructor. Under the heading of “scientific teaching,” Handelsman and her colleagues in a number of initiatives and publications have promoted the concept that science teaching should be directed by the same principles that inform scientific research activity. Specifically as developed by Handelsman et al. (2007) scientific teaching should involve the same levels of “critical thinking, rigor, creativity and the spirit of experimentation” as those used by scientists in their research. The concept of scientific teaching integrates the activity of conducting scientific research with the activity of scientific teaching. Importantly, the concept of scientific teaching rests upon two core principles: the need for engaging educational experiences based on an understanding of scientific inquiry and the need for evidence that will allow the evaluation of all educational activities.

The outcome of these three developments in science education is the proposition that science education should focus on scientific inquiry, should optimally take place within a laboratory setting, and be directed by the principles of scientific teaching. An educational program of this sort should engage students in the process of scientific inquiry and allow scientist-instructors the ability to measure and evaluate the quality and content of the education their students are receiving. It is this last point that the current book addresses. The aim of this book is to provide scientist-educators with a set of conceptual tools that will allow them to create an assessment strategy and assessment tools that are appropriate for the assessment of an educational, scientific inquiry program situated within a laboratory setting. This book builds upon the basic ideas of scientific teaching by providing an approach to the development of assessment tools that can be used in a variety of settings and thus allow scientists to consider carefully the quality and learning outcomes of their teaching. It should be noted that the current book addresses assessment within the confines of the educational developments specified above, namely teaching scientific inquiry through in-laboratory experiences. This is a relatively complex educational context and accordingly specific approaches to this distinctive educational setting are developed in this book.

1.4 Active Assessment Defined

Active assessment is the process through which scientist-educators develop an assessment strategy and assessment tools that provide significant information that enhances the active learning experience of students involved in the scientific inquiry process. Active assessment is a process in that the scientist-educator is actively involved in the development of their own assessment tools. In other words, the scientist-educator is an active part of the process of understanding how her/his educational programs are assessed and is in no way a passive recipient of standardized, externally created assessment tools. In addition, active assessment is based on the idea that the students who are engaged in scientific inquiry within a laboratory setting experience the assessment process as integral to their scientific inquiry process.
and as a source of input that informs their work and understanding as scientists. The active assessment process requires the scientist-educator to be deeply involved in the design of an educational experience that provides serious feedback and makes the scientific inquiry process conceptually meaningful and scientifically valuable for both the instructor and student-researcher. The characteristics of active assessment can be summarized as follows:

1. Active assessment elicits significant information that can be used to assess the quality and content of the educational inquiry program.
2. Active assessment is embedded within the scientific inquiry process and reflects meaningful practice within the laboratory setting.
3. Active assessment is developed by the scientist-educator and reflects the core procedural and substantive understandings of the specific scientific inquiry process that is utilized within the laboratory setting.
4. Active assessment is meaningful to the student-researcher and provides significant feedback that contributes to the educational and scientific development of the student-researcher.

1.5 The Underpinning Principles of Active Assessment

The idea of active assessment as defined above is based on several basic principles of educational practice. The first characteristic – *Active assessment elicits significant information that can be used to assess the quality and content of the educational inquiry program* – is tied to the core concept of all assessment research that educational processes and outcomes can and should be measured. From the perspective of scientific teaching, decisions in relation to educational practice should be based on the presence of relevant and comprehensive data concerning the process and outcomes of learning. Active assessment as an approach to assessment is designed to provide quality-contextualized information that in a very direct way measures those aspects of the scientific inquiry program that represent moments of meaningful knowledge development in the student-researchers’ projects. Evidence collected from the process of active assessment concerning students’ knowledge can be used to evaluate the state of learning within the program by pinpointing areas in which development has (or has not) occurred. This information can be used to modify or enhance various components of the educational program. In other words, active assessment should provide the evidence through which the educational value of the scientific inquiry program can be measured.

The second characteristic – *Active assessment is embedded within the scientific inquiry process and reflects meaningful practice within the laboratory setting* – addresses the core understanding that scientific inquiry is a particularized activity and that learning science is best advanced through contextualized understanding. The basic educational principle exemplified within the in-laboratory scientific inquiry process is termed situated learning (Lave & Wenger, 1991). Situated
learning proposes that meaningful learning takes place within authentic settings and as part of a process of participation within a community of practice. A series of researchers have posited the constructivist argument that learning is enhanced through participation in a contextualized, personally meaningful, problem-solving setting (Brown, Collins, & Duguid, 1989; Polman, 1999; Roth 1995; Wenger 1998; Williams & Hmelo, 1998). As argued by Lee and Songer (2003) authentic science practices within the cultural context of real scientific inquiry are important components of advancing scientific understanding. The movement of science education into the professional laboratory enhances the possibility that students get access to the personnel, tasks, and culture of scientific work. Active assessment extends this position on the importance of embedding learning within authentic contexts. As developed in this book, active assessment situates assessment within the framework and in relation to the authentic tasks and aspects of specific scientific inquiries. Rather than an imposed generalized approach, active assessment promotes contextualized assessment tasks tied to the specificity of the particular scientific inquiries. The assessment procedures closely reflect the actual work done in the laboratory and as such can provide evidence of understanding on the knowledge, concepts, and skills required to complete that particular scientific inquiry process. The embedded nature of the active assessment should make the procedures of assessment meaningful to the student-researchers as well as provide direct evidence of learning outcomes for the researcher-educator.

The third characteristic – *Active assessment is developed by the scientist-educator and reflects the core procedural and substantial understandings of the specific scientific inquiry process that is utilized within the laboratory setting* – addresses the central role the scientist-educators play in the development of both the educational program and its assessment procedures. The process of active assessment situates the scientist-educator in a central role in relation to the development of educational programs. Based on the principles developed within the framework of *Scientific Teaching*, the researcher-educator having developed an educational program should have interest in ways to measure knowledge acquired within that program. As described in the last point the types of task developed in the process of active assessment are closely related to the specific characteristics of the particular scientific inquiry process utilized as the basis for the educational program. The process of active assessment assumes that the specific scientific inquiry process is tied to the prime researchers’ own scientific research agenda. Accordingly, the prime researcher is best qualified to know what is significant from both a substantial and procedural perspective. It is these understandings of the significance of particular types of knowledge within the scientific inquiry process that ultimately constitute the basic components of the active assessment process. In this sense the scientist-educator utilizes her/his developed understanding of science to make decisions concerning what is of significance for assessment.

The fourth characteristic – *Active assessment is meaningful to the student-researcher and provides significant feedback that contributes to the educational and scientific development of the student-researcher* – addresses the role of metacognition as an important aspect of learning. As described by Baird (1990)
metacognition consists of the learner consciously recognizing aspects of their own learning and modifying or redirecting their educational actions in accordance with this understanding. Significant contextualized feedback can potentially play an important role in directing the students’ attention toward areas where further understanding is required. On a different level, active assessment can play a role in making sure that the educational laboratory experience is not just a case of manipulating procedures but rather a truly conceptual process. White and Gunstone (1992) posit that the integration of the laboratory experience with metacognitive tasks enhances the learning of science. As conceptualized within this book, the aims of active assessment are formative, diagnostic, and summative. The student-researcher’s usage of the assessment tools should provide information that is useful for the student in recognizing what has been learnt and what still needs to be learnt. The embedded, contextualized nature of the assessment process should ensure that the student gets specific feedback that is directly relevant to the scientific inquiry that is being followed. The outcomes of the assessment procedures need to become feedback to the student on the different aspects of the inquiry process that they are involved with. As a metacognitive process the feedback from the assessment procedures should direct the student to focus on areas that need to be addressed and thus allow learning to be individualized to the student’s specific needs. Meaningful assessment that provides direct feedback to the student-researcher is a central aspect of the active assessment process.

1.6 On the Importance of Active Engagement

As can be seen in the definitions of active assessment provided above, active engagement and problem solving on the part of the scientist-educator and the student-researcher are core aspects of the approach to assessment that is developed in this book. This assertion concerning the importance of the active engagement of both researcher-educator and student-researcher within the active assessment process needs some clarification. In relation to the scientist-educator the emphasis on active engagement with the assessment process has two aspects to it. First, scientist-educators are coming under increasing pressure to become involved with educational processes. This pressure is directed by national educational bodies, funding agencies, university administers, department chairs, and most importantly from a deep-rooted sense that able students are choosing other professions rather than the sciences. The outcome of these forces is that the research scientist is being redefined, at least in part, as an educator and producer of quality science education programs. Understandably many research scientists do not feel comfortable with this role and may even feel that this is an imposition on the more significant aspects of their scientific work. In addition many research scientists do not feel that they are really qualified to make educational decisions in particular in relation to the assessment of learning outcomes.
The solution posed within this book is to redress the supposed distance between educational and research activity and to situate active assessment within the context of authentic scientific inquiry. The idea as developed with the work of Hatfull et al. (2006) and Hanauer et al. (2006) is to integrate the educational scientific inquiry process within the scientific research agenda of the professional laboratory. In the second half of this book, one such program The PHIRE program (Phage Hunting Integrating Research and Education) situated at the Bacteriophage Institute of Pittsburgh at the University of Pittsburgh will be described in detail. The idea behind the PHIRE program and other programs like it is that students are involved with authentic scientific inquiry as full partners in the process. Naturally, the specifics of the educational research project need to be carefully chosen and not every research project conducted in a professional laboratory is compatible with educational processes (see Hatfull et al., 2006 for a full discussion of this point).

Of utmost importance to this approach to science education is the idea that the prime educator is also the prime researcher and that these two roles are compatible. The educational component of the laboratory’s research agenda is designed to serve the overall mission of the laboratory as a site of scientific knowledge construction. In an environment of this kind the scientist-educator’s engagement is obvious. In the same way as the head of a laboratory is fully committed and actively engaged with the research agendas of the laboratory, the researcher-educator is engaged with the educational process of the student-researchers. All serve the same aim of constructing scientific knowledge. Active involvement with the process of assessment design in this framework is an extension of the same engagement. As will be outlined in the following chapters of this book, the process of developing an assessment procedure is directly tied to the prime researcher’s understanding of the specific process of scientific inquiry that is being utilized in the educational program. In this sense, engagement in active assessment means really utilizing the prime researcher-educators’ understanding of the scientific inquiry processes of her/his laboratory that have become part of an educational program.

The second aspect of the importance of scientist-educator engagement with active assessment could be defined as a quality of life issue. The development of an assessment procedure by the prime laboratory researcher can provide both a sense of control and achievement in relation to the educational component of the inquiry process. When done well, active assessment provides very close, real-time data on the quality of the students’ knowledge, understanding, and work within the scientific inquiry process. Rather than thinking of assessment as an external process that is beyond the control of the prime researcher (and might influence future funding), as the designer of the assessment process, information concerning the development of student understanding is immediately accessible to the prime researcher-educator and, if necessary, changes to the program can be made to ensure that a quality education in the sciences is achieved. In addition, there is the potential for enormous satisfaction when evidence of student learning that is really meaningful is found. A by-product of this process of active assessment is the option for a second line of laboratory research. By becoming directly involved and engaged with the assessment process the scientist-educator develops a secondary research agenda that both allows
the educational program to evolve and improve as well as offering a new source of publishable research. Within the PHIRE program mentioned above, research is published on both the scientific and educational research that is conducted within the laboratory.

The argument for student engagement in the educational process is supported by some quite clear research outcomes and current understandings of quality educational practice. Handelsman et al. (2007) in their discussion of active learning review several studies that support the position that active engagement in scientific education enhances the retention of scientific information from courses. Knight and Wood (2005) utilizing problem-solving activities, discussions of journal articles, and learning assistants demonstrated increased knowledge retention. Similar results for the benefits of active engagement for science education have been presented by Udovic et al. (2002), Hake (1998), Beichner et al. (1999), and Ebert-May et al. (2003). Handelsman et al. (2007) summarizes their discussion by stating “Consistent across these research studies are the findings that active engagement enhances learning and retention, and that active learning builds higher-order thinking skills” (p. 26). The concept of active assessment is the attempt to make the assessment process a direct part of generating active engagement on the part of student-researcher in relation to the educational process that they are involved in. The underpinning concepts of embedding active assessment within the cultural practices of the laboratory and making the assessment tasks authentic and meaningful are designed to enhance the engagement of the student in the learning process and to provide real feedback to allow the student to play an active role in their own learning.

1.7 The Design of This Book

The aim of this book is to provide scientist-educators with an approach to the assessment of educational programs situated within laboratory settings and involving scientific inquiries. The book is designed as two sections. The first section of the book develops the theoretical and practical concerns involved in the process of active assessment. The aim of this section is to ensure that the reader is grounded in the literature, concepts, and practical aspects that inform the approach developed within this book. The second section provides a case study that exemplifies the approach developed within this book. Chapter 1 introduces the concept of active assessment and some of the sources and reasons for this approach. Chapter 2 involves a discussion of the concept of scientific inquiry. The direction of this chapter is to redefine scientific inquiry in contextual–operational terms that form the basis of the type of authentic assessment proposed in this book to be a possibility. Chapter 3 offers an introduction to assessment by presenting and explicating the basic concepts of assessment. Chapter 4 develops a specific approach to the assessment of scientific inquiry. The chapter reviews research that has been conducted on the assessment of scientific inquiry and then develops an approach to assessment that places high value of authenticity. Chapter 5 operationalizes the ideas of authentic, active assessment
and provides the reader with a conceptual framework that allows the development of an assessment program for educational scientific inquiry programs. The second section of this book provides a case study in which the ideas and frameworks presented in the first section are exemplified. Chapter 6 presents an introduction and overview of the educational scientific inquiry program (entitled the PHIRE – Phage Hunting Integrating Research and Education – program) in which the ideas for assessment presented in this book were developed and will be exemplified. Chapter 7 presents the specific assessment strategy that was developed for and utilized in this program. Chapter 8 presents a description of the specific tools developed for the PHIRE program assessment. Finally Chap. 9 is written from the perspective of the scientist-educator and presents the prime researcher’s experience of utilizing the process of active assessment within his program.
Chapter 2
Conceptualizing Scientific Inquiry

2.1 Introduction

In order to develop a strategy for the assessment of scientific inquiry in a laboratory setting, a theoretical construct of the components of scientific inquiry needs to be developed. A basic principle of any assessment procedure is that the starting point of any project is the specification of the object that is to be assessed. In the case of active assessment, the object to be assessed is scientific inquiry. Unfortunately, as discussed by Hofstein and Lunetta (2003) in their meta-analysis of 20 years of research concerning the use of the laboratory as a site for education, the term scientific inquiry has been described and defined in a variety of ways leading to the need for “greater precision and consistency” in the explanation of this term. The aim of the current chapter is to provide an understanding of some of the complexities within the educational context in defining scientific inquiry. The main problem that this chapter addresses is the definition of scientific inquiry as an object that can form the basis for the development of a program of assessment. The definition of scientific inquiry developed in this chapter posits a very significant role for the contextualized nature of scientific inquiry.

2.2 The Diversity of Scientific Inquiry

Since the widely referenced and acknowledged National Research Council (NRC, 1996) definition of National Science Education Standards there has been renewed recognition that the enhancement and propagation of scientific inquiry is one of the core elements of scientific education. The basic idea is that to learn science involves conducting activities that address the procedural and epistemological aspects of science. The National Science Education Standards (NRC, 1996) state that “scientific inquiry is at the heart of science and science learning” (p. 15). As conceptualized by the NRC, scientific inquiry includes a range of activities involved and related to the scientific process. Specifically the NRC defines inquiry in the following
Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations” (1996, p. 23).

However, while there is broad agreement over the potential significance of scientific inquiry for science education, the definition of scientific inquiry has to a certain extent been quite elusive and on a practical level difficult to implement. Hodson (1996), in a historical overview and educational critique of the scientific inquiry movement, deconstructs the simplistic notion of scientific inquiry as a decontextualized set of abstract principles (such as those posed in the NRC definition) that can easily be transferred from one scientific context to another. Hodson (1996) points out that actual scientific inquiry is infused with specific theoretical knowledge and hence contextualized in very specific ways. As stated by Hodson, “The difficulty of an observational task depends crucially on what is being observed and what constitutes appropriate or significant observation. In other words, the task is governed by the nature of the concepts involved” (1996, p. 126). Meaningful scientific inquiry is contextualized within a specific and developed knowledge structure and not the abstract application of procedural knowledge.

A different direction of critique of the scientific inquiry movement comes from studies of in-school manifestations of scientific inquiry. Millar (1998) is skeptical about the ability of in-school laboratory experiments to provide alternative understandings of accepted substantive descriptions of scientific concepts. Millar (1998) describes this type of scientific inquiry as a rhetorical form designed to manipulate results in order to provide specific, historically defined answers. Nott and Smith (1995) show how teachers manipulate the actual classroom demonstrations so that they conform to the accepted position on how they are supposed to perform. Hanauer (2006) in a study of elementary school students reveals how under the heading of scientific inquiry a multimodal structure of oral, written, visual, and physical forms of communication direct students to required and predefined results. All these studies propose that scientific inquiry within the context of pedagogical discourse can become a persuasive communicative tool designed to convince students of the correctness of predefined scientific concepts rather than a tool of scientific discovery.

Several studies have analyzed the handbooks (manuals) that are used in schools to direct scientific inquiry laboratories. In an early study using a classification tool termed The Laboratory Structure and Task Analysis Inventory (Tamir & Lunetta, 1978; Lunetta & Tamir, 1979), Tamir and Lunetta (1981) coded three high school curricula in the disciplines of biology, physics, and chemistry. Their findings found that “almost all investigations were highly structured” and that “Seldom, if ever, are students asked to: (a) formulate a question to be investigated; (b) formulate an hypothesis to be tested; (c) predict experimental results; work according to their own design; (d) formulate new questions based on the investigation” (Tamir & Lunetta, 1981, p. 482). In addition these researchers point out that students are “often asked
to perform a variety of manipulative and observational procedures and to interpret the results of their investigations” (Tamir and Lunetta, 1981, p. 482). In a later study using the same classification tool, Germann et al. (1996) studied nine biology laboratory manuals and found, once again, that biology laboratories are highly structured and that students were seldom provided with opportunities to “pose a question to be investigated; formulate a hypothesis to be tested; predict experimental results; design observation, measurement and experimental procedures; work according to their own design; or formulate a new question or apply an experimental technique based on the investigation they performed” (Germann, Haskins, & Auls, 1996, p. 493). The results of both these studies are strikingly similar and suggest that a wide range of in-school laboratory experiences emphasize physical manipulation over conceptualization and discovery.

A different approach to the definition of scientific inquiry was developed by Chinn and Malhotra (2002). In these researchers work, the epistemological and reasoning aspects of professional science were compared with school manifestations of scientific inquiry. Using the technique of examining school textbooks for hands-on activities, these researchers differentiate between three types of simple inquiry tasks: simple experiments (a single factor experimental design), simple observations (the careful observation and description of an object), and simple illustrations (following a specific procedure). These types of scientific inquiry were compared to authentic scientific inquiry in relation to the cognitive process involved and the epistemological aspects of the tasks. In relation to the cognitive processes of generating research question, designing studies, making observations, developing theories, and studying research reports, the differences between in-school scientific inquiry and authentic inquiry are pronounced. In relation to question generation and study design, in-school scientific inquiry is teacher-directed with students following directions, whereas, scientists function much more as independent problem solvers. In relation to making observations, explaining results, and developing theories, in-school scientific inquiry thought processes are directed toward straightforwardly addressing research questions without addressing the problems of observer bias, data transformation, experimental flaws, generalizability, theory development, conflicting data and inconsistencies, and more extensive literature. In authentic science the complications and ontological status of any scientific statement is a consistent concern.

In Chinn and Malhotra’s (2002) study the epistemological underpinnings of in-school scientific inquiry activities are differentiated from authentic scientific inquiry in relation to the purposes, nature of reasoning, and social construction of knowledge. Specifically authentic scientific inquiry is directed at the construction of knowledge through a variety of forms of argumentation and in the context of a community of researchers. In-school scientific inquiry does not develop knowledge, is limited in its forms of argumentation, and is not related to the wider community of scientific researchers. Zachos et al. (2000) extends this differentiation through the distinction between what they term “personal” and “cultural” discoveries. The discoveries of scientists have a historical and cultural aspect in that they constitute moments at which new knowledge is created. In this sense they are cultural discoveries. Personal discoveries involve the development of new knowledge for the
individual – a moment in which an understanding of phenomena is achieved – but not a move forward for the wider scientific community. In this sense they are personal discoveries. As seen in Chinn and Malhorta (2002), in-school scientific inquiry is designed to produce personal and not cultural discoveries.

As exemplified in the studies reviewed above, a major aspect of the confusion over the concept of scientific inquiry results from the fact that under the heading of scientific inquiry a wide range of different activities are being conducted. Wenning (2005; 2007) provides a useful heuristic through which this range of scientific inquiry activities can be addressed. Wenning (2005; 2007) develops a continuum of types of scientific inquiry that differentiates forms and educational uses of inquiry in relation to the degree of teacher control and required intellectual sophistication. Wenning (2005; 2007) presents the following levels and types of scientific inquiry:

1. **Discovery Learning**: This is the most basic form of scientific inquiry and consists of a teacher-controlled activity through which students are directed to make specific observations and reach predefined conclusions.

2. **Interactive Demonstrations**: This consists of a teacher-controlled manipulation of a scientific demonstration and the request for a prediction or the explanation of the phenomena. The teacher is in complete control of the demonstration, questions, and responses. The emphasis is on the teacher’s manipulation of scientific equipment.

3. **Inquiry Lessons**: This consists of a teacher-controlled demonstration of an experimental procedure. This demonstration of an experiment is accompanied by a verbalization of the conceptual and physical aspects of the experimental design. The teacher also asks leading questions and models the thought processes involved in scientific inquiry.

4. **Guided Inquiry Labs**: This consists of a teacher-directed student inquiry. Students conduct a scientific inquiry that is directed by a question presented by the teacher and lab procedures defined and guided by the teacher. Students are directed to find the answer to a specific question through the usage of a provided set of procedures (Herron, 1971).

5. **Bounded Inquiry Labs**: This consists of a student scientific inquiry that is directed by a question that is identified and posed by the teacher. The students are expected to design the experiment and conduct the scientific inquiry (Herron, 1971).

6. **Free Inquiry Labs**: This consists of a scientific inquiry that is directed by a question identified and posed by the student and an inquiry process designed and conducted by the student.

7. **Pure Hypothetical Inquiry**: This form of inquiry involves students developing hypothetical explanations of laws and explanations of physical phenomena based on empirical outcomes. This type of inquiry emphasizes pure hypothetical reasoning.

8. **Applied Hypothetical Inquiry**: This form of inquiry consists of problem-based learning in which a specific real-world problem is presented to the student and
through a process of hypothesis formulation on the basis of factual knowledge solutions and explanations are posed. Posed solutions are supported through logical argumentation and informed reasoning.

Wenning’s (2005; 2007) conceptual scheme of types of scientific inquiry, building upon Herron’s (1971) earlier work, clarifies the problem with the definition of scientific inquiry. Scientific inquiry is an umbrella term for a range of educational and professional activities within the sciences. While the rhetoric has tended to conflate scientific inquiry with authentic professional science, the majority of studies that consider in-school activity suggest that there are significant conceptual, physical, and epistemological differences between professional and educational scientific inquiry.

2.3 The Characteristics of Authentic Scientific Inquiry

The discussion in the previous section has brought us forward in the sense that it is clear that scientific inquiry covers a range of types of activity that are very different from what is usually done within professional science. However, this discussion does not solve the main problem that this chapter addresses – the definition of scientific inquiry as an object that can form the basis for the development of a program of assessment. The context of scientific inquiry that this book addresses is different from the ones that were discussed in the previous section. This book describes an approach to the development of assessment procedures for in-laboratory, educational scientific inquiry programs designed around authentic scientific research questions, directed by a real research agenda of interest to the wider scientific community and coordinated by an active research scientist. This context presents from an educational perspective a new configuration of the components of scientific inquiry. Using the concepts developed in the previous section, the description of this form of scientific inquiry can be characterized as follows:

1. Development of Personal and Cultural Knowledge: A central aspect of the type of in-laboratory educational scientific inquiry programs addressed in this book is that they are directed by the presence of an active scientist with an authentic, scientifically valuable research agenda. In other words, the research that is being conducted by the student-researcher is designed to provide both personal and cultural knowledge. The scientific inquiry process may be educational in that it provides the student with the experience of learning how research is conducted, but ultimately the aim of this research is not limited to the realm of education but rather has the goal of the creation of new scientific knowledge that is publishable.

2. Contextualized Scientific Knowledge: The development of a research agenda and the ability to identify what research needs to be conducted and is of value for the wider scientific community require a comprehensive and sophisticated understanding of the scientific knowledge that exists within the specific discipline. By definition, research that is conducted with the aim of producing cultural knowledge is contextualized within a specific knowledge structure. Accordingly,
students studying within a framework of this kind are exposed to contextualized scientific inquiry practices relevant to a specific body of knowledge and are not involved in the application of abstract concepts of scientific inquiry. In addition, the assumption of a program of this sort is that the knowledge that will be produced will be presented in the settings and formats that characterize scientific communication such as research articles, research reports, posters, and conference presentations.

3. The Progression Toward High-Order Problem Solving: The cognitive aspect of an educational scientific inquiry program that is designed to produce publishable knowledge involves the movement toward argumentation and problem solving that is used within professional science. In other words, all the complexities of real research, such as the coordination of data and theory, the design of research, the resolution of problems that occur, and the discussion of anomalous results, are part of the educational program. As argued by Hatfull et al. (2006) for this to be practical the educational scientific inquiry program has to have an accessible entry point and progressively move to more complex structures. This creates a scientific inquiry program that starts from a research process that is guided and directed by knowledge, questions, and procedures from the instructor to a situation in which the student-researcher works independently in coordination and discussion with other members of the laboratory. In the terms developed by Herron (1971) and Wenning (2005) this involves the movement from a guided inquiry laboratory to a free inquiry laboratory. It is important to note that the last stages of the research project are open ended and cannot be predicted. They evolve as a result of informed decisions made as the situations arise. One ramification of this progression is that the student cannot stay on the level of physical manipulation of laboratory equipment but rather must acquire the substantive and procedural knowledge relevant to the discipline within which the specific scientific inquiry is being conducted.

4. Social Interaction for Scientific Goals: Professional science is characterized by extensive social interaction with other scientists. Science is not done alone but rather is the result of a community. In educational, scientific inquiry programs designed to produce cultural knowledge the wider community of scientists within the laboratory and beyond the laboratory will be addressed. In the early stages of the educational program the scientific inquiry process is guided and directed which involves extensive interaction with professional faculty in order to facilitate the development of relevant knowledge; at later stages of the process there is also extensive interaction with professional scientists but at this stage this consists of working out how the research agenda of the whole laboratory can be moved forward. Within programs of this type, student-researchers who create cultural knowledge are expected to interact with the other scientists through the professional lines of communication such as professional conferences.

5. Scientific Inquiry as a Multi-stage and Multi-representational Process: A professional scientific research agenda is characterized by the presence of multiple stages of research over an extended period of time. A scientific inquiry progresses through a series of laboratory stages which often involve the
manipulation of various properties of the physical world. At each stage different physical outcomes are found and recorded in writing and visual formats. As described by Latour and Woolgar (1986), actions taken within the laboratory become meaningful when they are transformed into representational inscriptions. In this sense a lot of science is visual and representational. Accordingly, a scientific research process can be described as a series of stages each characterized by the development of a specific representational outcome.

2.4 An Analytical Framework for the Definition of Scientific Inquiry

Based on these characteristics of the educational scientific inquiry program presented above, a two-part analytical frame can be proposed to define scientific inquiry. However, before the model is described it should be remembered that the basic principle of this model is that this model is fully contextualized within the framework of a professional laboratory involved in an educational scientific inquiry designed to produce cultural knowledge. Accordingly, it is assumed that the actual definition of a scientific inquiry that will form the basis for the development of an assessment program results from the specific application of the proposed analytical framework to the definition of the specific scientific inquiry process that forms the basis for the educational program. The proposed model has two analytical parts: it proposes a series of types of knowledge that are significant for any scientific inquiry process and it then proposes an organizational structure comprised of two intersecting axes for these knowledge types.

In a study that was designed to produce assessment tools, Hanauer (2007) analyzed the scientific process of a particular educational in-laboratory scientific inquiry program from a multimodal information processing perspective (Hanauer, 2006). As presented by Hanauer (2007) the analysis of this scientific inquiry process involved four different types of information: physical knowledge, representational knowledge, cognitive knowledge, and presentational knowledge. These different types of knowledge are assumed to work together in a multimodal-layered construct and were each considered to be knowledge that needs to be assessed when considering whether scientific inquiry knowledge is being acquired. These knowledge types were defined as follows:

- **Physical knowledge** consists of knowledge required to actually perform the laboratory tasks involved in scientific inquiry.
- **Representational knowledge** consists of the written and visual representations used within the laboratory.
- **Cognitive knowledge** consists of background disciplinary knowledge of scientific content and thinking abilities such as problem solving, decision making, and calculation.
• *Presentational knowledge* consists of the ability to summarize understandings from research, to conceptualize these in manner that is valuable for the scientific community, and present them in the formats that are used by the scientific community.

The approach taken here is that in order to understand and operationally define the process of scientific inquiry, in addition to an analysis of the stages of a scientific process, scientific inquiry needs to be analyzed in terms of the types of knowledge that are utilized as part of the inquiry process. This type of analysis produces a description of the scientific inquiry process and allows assessment materials to be developed on this basis. The types of knowledge proposed by Hanauer (2007) are modeled in Fig. 2.1.

As represented in Fig. 2.1, the process of scientific inquiry is a multimodal and multilayered phenomenon that integrates four different types of knowledge. Cognitive knowledge provides the background scientific information that is crucial to the actual understanding of the science involved in the scientific inquiry process. Cognitive knowledge also involves the ability to make informed decisions, interpret physical and visual results, and make calculations. This information source feeds into the physical activities that are actually conducted in the wet laboratory. The results of the laboratory process are in the form of visual and written representations that are stored in most cases within the framework of the laboratory notebook and as physical outputs from the laboratory work itself (such as plates or gels). To understand the visual representations that were produced, cognitive knowledge is

![Fig. 2.1](image_url) A Schematic Representation of the Types of Knowledge Involved in a Scientific Inquiry Process
applied and decisions are made as to the future directions that need to be taken. Finally, presentational knowledge is to be considered the ability to produce a product that is useful for the wider scientific community to gain access to the research that was conducted. Presentational knowledge consists of the public presentation of conference posters (or perhaps written research papers or oral presentations) at professional conferences or within educational settings. The production of a poster integrates knowledge that comes from the notebook entries and other visual representations. The actual creation of a poster (or written/oral paper) involves the summarization and reconceptualization of the actual research that was conducted and the decision as to what makes this information important for the wider scientific community. As such to create a poster requires both the representational and cognitive knowledge sources used during the scientific inquiry. It should be noted that all the knowledge types presented above are of significance within the educational scientific inquiry program. It is the interaction between the different knowledge types that creates some of the complexity in assessing and understanding the scientific inquiry process.

The development of an assessment strategy for scientific inquiry requires a basic description of the components of scientific inquiry and the way these components are organized. The organizational approach used here specifies that scientific inquiry can be defined by two theoretical axes: the axis of knowledge source and the axis of stages of a specific scientific inquiry process. The basic idea behind this organizational structure is that scientific inquiry is a multi-stage process that involves the development of a series of in-lab outcomes (representations) over an extended period of time. Accordingly the definition of a scientific inquiry process consists of understanding the specific aspects of each knowledge source and how they develop over a period of time. Another way of understanding this model is to think of an extended scientific inquiry as a collection of much smaller scientific inquiries each of which produces a definable outcome and that collectively develop toward a research finding. For each specific stage, cognitive, physical, and representational knowledge is applied in order to reach the desired outcome, so that at different stages of the scientific inquiry different knowledge is required. Accordingly as the student progresses through the educational program knowledge will develop in relation to all the types of knowledge. From an assessment perspective, as will be discussed in later chapters, in the process of active assessment each of these axes requires a detailed analysis by the prime scientist-educator of the specific research project that is being utilized. Together the definition of knowledge for each of these two axes provides a detailed and operative description of a scientific process that can be used to generate a comprehensive assessment strategy. Figure 2.2 presents a schematic representation of the scientific inquiry process.

As an operational tool the two axes provide active researchers or observers of a scientific inquiry process a terminology and a conceptual heuristic with which to describe a scientific project in terms that are useful for assessment and educational design. The particular aspects and operational definitions of this framework are presented in Chap. 5. As seen in Fig. 2.2 the first axis divides a scientific inquiry process into representational “milestones.” The idea is that a scientific process can
be divided into stages according to the laboratory products that need to be produced. These products are termed representations in that they fulfill a representational role as the actual outcome of both a thinking process and physical laboratory activities. By analyzing the first axis a series of stages with specific outcomes can be defined and the overall process of scientific inquiry can be explicated. The second axis considers the types of knowledge that are required in order to produce the required laboratory representational “milestones”. Every laboratory product results from the application of cognitive, representational, and physical knowledge sources. The analysis that needs to be done here is of exactly what knowledge is required to produce and comprehend each representation. This knowledge should be defined under the heading of cognitive, physical, and representational knowledge types. If both axes 1 and 2 are analyzed a detailed and operationally functional description of a specific scientific process should emerge. The detailed description of a specific scientific inquiry process is the basis upon which a comprehensive assessment program can be designed. In broad terms every assessment of an inquiry process would want to address the four knowledge sources of cognitive, physical, representational, and presentational knowledge and do so in relation to the different stages of the process itself.

### 2.5 Chapter Summary

The aim of this chapter was to provide an understanding of the concept of scientific inquiry that is applicable as a basis for the development of an assessment program. The following ideas and concepts were defined:
• The term scientific inquiry as manifest in different educational settings covers a wide range of diverse activities.
• The differences in types of scientific inquiry can be organized along a continuum according to the degree of teacher control and intellectual sophistication involved in each type of inquiry.
• Types of scientific inquiry can also be defined according to whether they produce cultural knowledge or personal knowledge.
• Authentic scientific inquiry is defined according to five characteristics: development of personal and cultural knowledge; contextualized scientific knowledge; the progression toward high-order problem solving; social interaction for scientific goals; and scientific inquiry as a multi-stage and multi-representational process.
• The definition of scientific inquiry that forms the basis for the development of an assessment program consists of a two-part analytical frame: the definition of knowledge types relevant to scientific inquiry and the definition of an organizational frame for these knowledge types.
• Four types of knowledge are significant for the definition of a specific scientific inquiry program: cognitive knowledge, physical knowledge, representational knowledge, and presentational knowledge. All four of these knowledge types are considered significant.
• These four types of knowledge are organized in a framework that consists of two intersecting axes: the axis of knowledge types and the axis of stages of a specific scientific inquiry. This framework describes scientific inquiry as multi-stage process that involves the development of a series of in-lab outcomes (representations) over an extended period of time.
• The definition of a scientific inquiry is contextualized within the framework of a specific research agenda in a particular scientific field that is directed by an active researcher.
Chapter 3
An Introduction to Assessment

3.1 Assessment and Science

Assessment like science is a practical, systematic, evidence-driven activity. Assessment like science produces knowledge that is used to make decisions about the nature and characteristics of the world. Simply put, assessment and science are conceptually compatible activities. It is therefore surprising that many scientists feel uncomfortable with the idea and practice of educational assessment within their educational activities. The aim of this chapter is to introduce some basic concepts and approaches to the development of assessment. As with other scientific activities, enhanced understanding of the underlying concepts and procedures that direct assessment should allow scientists an increased ability to understand their educational programs and practices. In addition, this should enhance the understanding of the specific approach to in-laboratory assessment that is developed in this book.

3.2 Assessment, Evaluation, and Testing

For many scientists one of the core points of confusion relates to the terminology of educational assessment. In this section, we will try and make some key distinctions clear. Zachos (2004) states that assessment “is the process of obtaining evidence to support inferences concerning the attainment of learning objectives” (p. 748). For Zachos, as for the authors of this book, assessment can be seen as a form of scientific inquiry process in which you collect data in order to understand student learning. Assessment in this sense is that process through which this information is attained and conclusions are reached through logical, reasoned extrapolation from collected data.

Assessment needs to be differentiated from the processes of evaluation and testing which are often used interchangeably. Evaluation refers to the process of using assessment data to judge the quality of a program or group of students. Testing refers to the tools and procedures used to elicit information that is part of the assessment
process. This book deals with internal evaluation of in-laboratory science education programs and provides tools which can be used for evaluation purposes. However, it does not deal directly with the relationship of evaluation to external educational policy or funding. Testing is dealt with as part of the process of collecting data utilized in the assessment process. The main focus of our discussion is to provide a framework for the systematic development of an assessment strategy that will allow scientists involved in scientific inquiry teaching to assess the learning of their students.

3.3 Summative, Formative, and Diagnostic Assessment

Within the existing assessment literature a series of different aims for assessment have been defined. The National Research Committee on Classroom Assessment and the National Science Education Standards differentiate between formative and summative assessment aims. The core difference between summative and formative assessment is not the tool used to collect the data but rather the function of the conclusions from the data. Summative assessment is used for the aim of “offering a cumulative summary of achievement level” (NRC, 2001, p. 5). Formative assessment is used to “inform teaching and/or to influence learning” (NRC, 2001, p. 5). Within educational contexts summative assessment data are used to provide final grading or placement for a student taking a class. Formative assessment is used by the instructor as input in order to redirect and improve classroom teaching. In this sense formative assessment is used as a tool that informs instruction. The NRC (2001) propose that formative assessment is directed by the question “Where are you now” and that the answer to this question allows the teacher to design educational interventions that will help the student to achieve the end goals of the course. The concept directing this understanding of formative assessment involves a feedback loop between the formative assessment data that have been collected and the specific educational decisions made by the science instructor.

There is a third aim for assessment that is worth addressing. Diagnostic assessment aims to provide the instructor and the student with information concerning the current understanding and performance of the student. Diagnostic assessment has the function of identifying the student’s specific strengths and weaknesses in relation to the information and procedures that are being learnt. The three aims of assessment discussed here and differentiated in relation to their functions can be summarized as follows:

- **Summative Assessment**: Aims to provide information for decisions concerning the final status of knowledge of the student
- **Formative Assessment**: Aims to provide information for decisions concerning the way to achieve the educational goals of the program
- **Diagnostic Assessment**: Aims to provide information for decisions on the current state of knowledge of the student.
These three aims for assessment should be seen as integrated and to a large extent utilizing the same data collection methods. Accordingly, an assessment strategy for an educational program usually involves all three of these assessment functions.

### 3.4 Distinctions in Elicitation Methods

There are a wide range of tools and procedures that can be used to elicit information useful for the assessment of student learning. A basic characteristic of all elicitation tasks is that they require students to respond to a particular prompt. The differences among elicitation tasks relate to

1. The modality of response and prompt (written, visual, physical, and/or oral)
2. The relationship of the task to authentic scientific practice (direct or indirect)
3. The discrete or integrative nature of the knowledge tested
4. The requirement to perform functions or provide knowledge in the response
5. Degree of provided structure

Elicitation tasks can be in any modality. It is common for elicitation tasks to have a written component in both the prompt and the response. But this is in no way a requirement. With a laboratory setting, for example, elicitation prompts could be verbal (for example, “show me how you use a micropipette”) and the response may be physical (for example, actual usage by a student of a micropipette). This aspect of elicitation items is termed the modality of the task.

The task that is being tested could be directly drawn from the real task used within science or it could be a test of one abstract skill assumed to be significant within science. When an elicitation task is directly drawn from a real task used in science, it is considered to be an authentic task. Its authenticity is defined by the closeness of the elicitation task to the characteristics and components of the real scientific task. In some testing frameworks indirect elicitation tasks are used. An indirect task elicits information concerning a specific abstracted aspect of science. The indirect task is not modeled on a real task in science but rather is the result of an analysis of what is significant in this specific discipline and level of science.

As well known by any scientist, actually conducting scientific inquiries requires the integration of a range of skills and knowledge. Authentic elicitation tasks that are modeled on real science tend to integrate a variety of different skills and knowledge bases. Authentic elicitation tasks tend to be integrative tasks. Not all elicitation tasks are integrative. An elicitation task can address an individual part of a much larger knowledge structure. For example a student could be asked about her/his knowledge concerning the characteristics of the cell structure of a microorganism without being asked to integrate this knowledge with anything else. When the knowledge elicited from a student is done in a way that presents this knowledge as an independent item, this is termed a discrete item of knowledge. Typically, discrete elicitation tasks test one or other independent skill or content knowledge that is considered an important
part of the educational aims of the course. Authentic tasks tend to be integrative and indirect tasks tend to consist of discrete items.

The prompt of an elicitation task can request that the student perform a task usually involving a form of problem solving. Tasks of this kind are termed performance tasks and the data elicited from a task of this kind is the ability of the student to do the task itself. Performance tasks address procedural knowledge of science. A prompt could also require the presentation of knowledge. Tasks of this kind are termed knowledge tasks and the data elicited from a task of this kind is the content of the knowledge presented. These tasks address substantive or content-based knowledge of science. There is a certain degree of knowledge and performance in every elicitation task. But there is difference in the intent of different types of elicitation tasks on this criterion. A performance task requires the student to actually do the task whereas a knowledge task request the presentation of knowledge.

A final characteristic of tasks refers to the varying degrees of provided structure. The structure of a task refers to the amount of guidance and direction that is provided to the student. For example, a scientific inquiry task could be open ended without any clarification of the specific tasks to be performed or it could consist of a series of prescribed directions that need to be performed. The former task would be considered to be relatively unstructured placing a large conceptual and planning responsibility on the student; whereas the latter task would be considered to be a highly structured task with detailed instructions.

For the purpose of assessment development, it can be helpful to differentiate between the specific manifestations of testing (such as multiple choice items) and the theoretical concepts that underpin these test types. Taken together the distinctions presented above in relation to elicitation tasks provide a way of classifying as well as generating elicitation methods. For example, consider a multiple choice test of substantive knowledge of a specific organism. A test of this kind would be written (modality), indirect (authenticity), discrete, knowledge-based, and structured. A test in which students are required to annotate a short sequence of DNA would be written, authentic, integrative, performance-based, and moderately structured.

3.5 Developing an Assessment Program

Tests of knowledge do not just happen. Like other forms of scientific inquiry, a basic principle of all educational assessment is that the construction of an assessment program is the result of careful and purposeful thinking. Doran et al. (2002) in their discussion of laboratory assessment specify eight stages that characterize the development of new assessments: State the Purpose, Select an Appropriate Task Format, Write or Modify the Task, Clarify Administrative Procedures, Develop a Scoring Rubric, Trial Test the Task, Revise the Task, and Analyze the Results. As specified by these researchers these stages function in a cyclical manner so that the development of appropriate assessment tasks is an on-going process of modification. These stages of development are defined as follows:
1. **State the Purpose:** The first stage of assessment development is conceptual. The initial stage of assessment development consists of defining the aims, functions, and types of knowledge to be assessed. For any assessment program there can be different aims for assessment including self-evaluation for improvement of the educational program, provision of data for external evaluation purposes, and the provision of feedback to students to enhance their educational process. The aims of program assessment should be tied to the actual aims of the educational program. If the program aims for generating knowledge of scientific facts then the assessment program should be assessing whether these facts have actually been retained by students. These real world aims are directly connected to the functions of assessment introduced in a previous section of this chapter. The concepts of summative, formative, and diagnostic assessment all define different potential functions of the assessment program. Having decided on the aims and function of the assessment program the types of knowledge to be assessed are addressed. Depending on the setting and aims of the educational program, assessment could address procedural knowledge or substantive, content knowledge of science. Furthermore the assessment designers at this stage need to consider the specific aspects of the types of knowledge that they wish to assess. Since the collection of assessment data is an extended process decisions have to be made as what will be chosen for assessment. It is assumed that all choices will address significant rather than trivial scientific knowledge.

2. **Select an Appropriate Format:** Having decided on the aims, functions, and specific knowledge that it is important to assess, the next stage of the planning process is to select an appropriate form of elicitation task for the assessment of this knowledge. As specified in a previous section of this chapter, types of elicitation task are defined according to five characteristics: modality, authenticity, discreteness, type of knowledge assessed (procedural or substantive), and degree of provided structure. The choice of specific elicitation task that will be used is directly connected to the decisions that have been made as to what is important to assess and what the functions of this data will be. For example, if the educational program aims to develop knowledge of a particular scientific concept, a written test of explicit knowledge may be suitable (the characteristics of written, indirect, discrete, and substantive). If an educational program aims to develop knowledge of a particular scientific procedure, the performance of the procedure within the laboratory and observation of the task may be suitable (the characteristics of physical, authentic, integrative, and procedural).

3. **Write or Modify the Task:** During this stage of the assessment development process, the planning that was conducted in the previous stages is translated into specific elicitation tasks that will be used for assessment purposes. Doran et al. (2002) state that issues of diversity need to be addressed at this stage. These include making the written tasks appropriate in style and language for different age groups, students with different language abilities, and students with a range of physical challenges. Before embarking on the actual writing and definition of an elicitation task it is worth considering the specific aims of the task that is being written. What knowledge does it address, is this knowledge significant,
and what type of knowledge is involved? The actual writing of an elicitation task involves making sure that what is required of the student does indeed address what the program wishes to assess. In some cases this can be straightforward as in the writing of an open-ended test question concerning factual scientific knowledge. In some cases, the development of an assessment task is more creative. It is well known within the context of testing that every elicitation task does indeed interfere with elicitation of student knowledge. Technically this is termed “task-effect” and consists of the negative contribution of the actual written tasks to the performance of the student on the task. As a result of the presence of task-effects, the writing of elicitation tasks needs careful consideration and piloting. The piloting and constant reevaluation of written tasks allows assessment developers to modify and change tasks when it is deemed necessary. In this sense the process of developing assessment tasks is an on-going task.

4. **Clarify Administrative Procedures**: Once a set of elicitation tasks has been written thought needs to be directed at the administrative and managerial aspects of the assessment process. The assessment process needs to be practical within the parameters of the educational process. Some of the practical aspects of assessment that need to be addressed are time constraints, safety procedures, ease of data collection, accessibility, and clarity. The completion of assessment procedures should be short, safe, accessible and clear to the students. While these aspects of an assessment process are obvious they are often forgotten as part of the planning of an assessment procedure. The importance of this stage is the idea that planning does not end with the development of the actual tools of elicitation but rather the actual process of administration of an assessment procedure requires consideration.

5. **Develop a Scoring Rubric**: The development of a scoring rubric consists of the exact specification of the way responses elicited from the student in relation to particular tasks will be understood by the assessment team. The development of an understanding of how the task will be completed by students is crucial for the construction of a meaningful assessment procedure. The scoring rubric should consist of sample answers, the assessment value attached to this type of answer, and the type of inference relating to the students’ understanding that can be inferred from this response. In addition the scoring rubric should specify exactly the type and content of knowledge that is expected from the elicitation task. This collection of conceptual tool should allow assessment developers to define how specific answers will be understood and what they say about student learning.

6. **Trial Test the Task**: Piloting of a developed assessment procedure consists of an empirical study of the quality of the tools that have been produced. Before using the developed assessment procedure on an extensive basis with your target population, it is a good idea to evaluate the quality and characteristics of the assessment procedure by piloting the assessment procedure. The piloting of assessment materials should mimic the actual proposed usage of the tasks within the educational setting. As a piloting procedure the conclusions should be used for the evaluation of the tools and the assessment procedures themselves.
7. *Revise the Task*: The development of an assessment procedure is an on-going activity. Following the trail test of the developed assessment tasks, conclusions concerning the quality of the tasks need to be reached and any required changes need to be made. This process of revising the tasks should be considered after every administration of the assessment procedures. Through the interactive description of assessment development presented here and its associated revision of written tasks, it is assumed that assessment procedures will improve over time and become an ever more reliable method of assessment.

8. *Analyze the Results*: In principle, the analysis of the results of the assessment procedure is based on the rubric that was developed for this purpose. However, as a form of scientific inquiry, it is worth considering all results also for the ways in which they diverge from the expected responses defined by the rubric. This aspect of considering the unexpected can lead to new insights into the nature of student understanding (or misunderstanding), comprehensiveness of the rating rubric, and aspects of the task itself. While the main aim of result analysis is to provide information that can be used to assess the state of student learning and program outcomes, the analysis of the result should also be seen as having the option of providing new insights into the educational process that might not have been foreseen before administration of the assessment procedure.

### 3.6 Chapter Summary

The aim of this chapter was to develop basic concepts relating to the process of assessment. The following concepts and ideas were defined:

- Assessment is a form of scientific inquiry process in which you collect data in order to understand student learning.
- Assessment is differentiated from evaluation (the judgmental usage of assessment data) and testing (the procedure for collecting assessment data).
- The functions of assessment can be summative, formative, or diagnostic.
- Elicitation tasks can be defined according to five characteristics: modality, authenticity, discreteness, type of knowledge assessed, and degree of structure.
- The process of developing an assessment procedure involves the stages of stating the purpose, selecting an appropriate task format, writing or modifying the task, clarifying administrative procedures, developing a scoring rubric, trial test of the task, revising the task, and analyzing the results.
Chapter 4
Assessing Scientific Inquiry

4.1 Introduction

The aim of the last chapter was to introduce some basic concepts and approaches to the development of assessment. In the current chapter, this discussion will be expanded so that it relates specifically to the issue of assessment that is addressed in this book – the assessment of scientific inquiry. This chapter reviews some of the proposals concerning scientific inquiry assessment and proposes its own approach to this issue. As discussed in previous chapters the approach taken to assessment in this book is one that places high significance on the contextual nature of scientific inquiry within an authentic research setting. Accordingly the chapter explores the concept of authentic performance assessment and discusses the importance of a deep understanding of the specific scientific inquiry process that is used in the educational program, as an initial stage in assessment development.

4.2 Approaches to the Assessment of Scientific Inquiry

The question of how to assess scientific inquiry has been addressed by several researchers and a series of assessment proposals have been presented. The differences among the approaches to assessment result from decisions concerning two issues: (1) the definition of the concept of scientific inquiry and (2) the type of elicitation method used. As described in the previous chapter the first stage in a process of developing an assessment program consists of the definition of the construct that one wishes to assess. As seen in the analysis of scientific inquiry in Chapter 2, a specific problem faced by anyone who wishes to create an assessment program for scientific inquiry is that scientific inquiry is a term that has a range of different meanings. To a certain extent the differences among the various approaches to the assessment of scientific inquiry result from the answer that each researcher provides to this question – what is scientific inquiry? A shared tendency among all the approaches to the scientific inquiry assessment is to define scientific
### Table 4.1 Approaches to the assessment of scientific inquiry

<table>
<thead>
<tr>
<th>Source of approach</th>
<th>Definition of scientific inquiry</th>
<th>Methods of elicitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zachos (2004); Zachos et al. (2000)</td>
<td><strong>Core capabilities of scientific inquiry:</strong> coordinating theories, underlying principles, precision, identifying error in measurement, and proportional reasoning</td>
<td>Written (open response), indirect, discrete, performance, and structured</td>
</tr>
<tr>
<td>Wenning (2007)</td>
<td><strong>Stages of scientific inquiry:</strong> identify a problem, formulate a hypothesis, generate a prediction, design an experiment, conduct an experiment, collect (organize and analyze) data, apply statistical methods and reach conclusions, explain unexpected results, and report results</td>
<td>Written (multiple choice), indirect, discrete, knowledge-based, and structured</td>
</tr>
<tr>
<td>Bryce et al. (1983); Davis (1989); Welford et al. (1985)</td>
<td><strong>Practical skills:</strong> planning, measurement, use of instruments and observation</td>
<td>Written, visual and physical, indirect, discrete, performance, and structured</td>
</tr>
<tr>
<td>Gott and Duggan (2002); Roberts and Gott (2003)</td>
<td><strong>Practical skills and concepts of evidence:</strong> instruments (relationships, calibration, and error), reliability and validity of measurement, choice of instrument, sampling, design (variable structure, validity, choosing values, accuracy, precision, and tables), reliability and validity of design, data presentation, statistical treatment of data, patterns in data, and reliability and validity of investigation</td>
<td>Written, discrete, indirect, knowledge, and structured</td>
</tr>
<tr>
<td>Shavelson, Solano-Flores, and Araceli Ruiz Primo (1998); Solano-Flores and Shavelson (1997); Solano-Flores et al. (1999)</td>
<td><strong>Science performance assessment:</strong> Four types of scientific inquiry – comparative, component investigation, classification, and observation</td>
<td>Physical and written, integrative, indirect, performance, and semi-structured</td>
</tr>
<tr>
<td>Lunsford and Melear (2004)</td>
<td><strong>Science as literacy act:</strong> Scientific inquiry defined through the products of the portfolios, research reports, student demonstrations, student journals, and concept maps</td>
<td>Written and physical, integrative, direct, knowledge, and structured</td>
</tr>
</tbody>
</table>

inquiry in relation to a series of abstracted elements. The differences concern the type of abstraction that is proposed as significant and the way these abstracted categories of scientific inquiry are operationalized through a method of data elicitation. Table 4.1 summarizes the approaches to the assessment of scientific inquiry reviewed here and the nature of the differences in relation to the definition of scientific inquiry and methods of elicitation.
Zachos (2004) and Zachos et al. (2000), as a first stage toward the development of a scientific inquiry assessment program, conducted a study designed specifically to find those core capabilities that characterize scientific inquiry. The results of their study propose the following list of what Zachos (2000) terms critical “scientific inquiry capabilities:

- Coordinating theories with evidence,
- Search for an underlying principle,
- Concern for precision,
- Identifying sources of error in measurement, and
- Proportional reasoning.” (p. 751)

In Zachos et al. (2000), these core capabilities were used as the conceptual basis for the design of a series of structured performance tasks. Three tasks – Floating and Sinking, Equilibrium on the Balance Beam and the Period of the Pendulum – were chosen from the research conducted by Inhelder and Piaget (1958) because they posed conceptual problems for students as they tried to construct and evaluate concepts of scientific phenomena. These tasks were considered to be useful for both developing and assessing scientific inquiry capabilities. However, Zachos (2004) suggests that the use of structured inquiry tasks and direct observation of performance is not feasible within educational systems in which the ratio of student-instructor is high and thus proposes that students be presented with the structured performance task and then asked to record “their responses on a special form” (p. 753). The responses would be assessed according to the core set of scientific inquiry capabilities that was defined by Zachos et al. (2000). In this sense the scientific inquiry capabilities are transformed into a set of performance indicators that are used to assess written responses to a structured inquiry task.

Wenning (2007), working within the context of physics education, designed a diagnostic multiple choice test of knowledge relevant for scientific inquiry. As specified by this test’s creator, this test is not designed to “authentically assess student abilities to conduct scientific inquiry” (Wenning, 2007, p. 24). The test is designed to provide measures of students’ development of knowledge concerning scientific inquiry and would be ideal for the pre- and post-testing of the development of such knowledge. As an initial stage of the test development process, Wenning (2005) conducted a literature review of academic materials that define scientific inquiry and proposed a list of stages that define the scientific inquiry process. This list was then reviewed by physics-teaching majors, scientists, and educators. The final list of stages was as follows:

- Identify a problem to be investigated.
- Formulate a hypothesis or model using induction from logic and evidence.
- Generate a prediction using deduction.
- Design an experimental procedure to test prediction.
- Conduct a scientific experiment, observation, or simulation to test the hypothesis or model.
- Collect meaningful data, organize, and analyze data accurately and precisely.
• Apply numerical and statistical methods to reach conclusions.
• Explain unexpected results.
• Use available technology to report, display, and defend results.

This list of the stages of scientific inquiry was used as a basis for the development of 40 multiple choice questions that cover the defined stages. Following repeated piloting, five questions were deleted and some questions changed and reworded leaving the final version of the test with 35 questions.

A different approach to the definition of scientific inquiry consisted of considering the practical, procedural aspects of scientific inquiry as central (Archenhold et al., 1988; Bryce et al., 1983; Davis, 1989; Welford et al., 1985). This type of approach sees issues such as measurement and instrument use as what needs to be taught and assessed within an educational inquiry program. The assessment of knowledge relating to practical skills could be in both a performance format in which students make readings of present instruments and a written format in which pictures of instruments are presented and students are required to provide a written response.

Gott and Duggan (2002) and Roberts and Gott (2003) working in the field of biology developed an approach to the assessment of scientific inquiry based on their concept of knowledge-based enquiry. According to Roberts and Gott (2003) knowledge-based enquiry is characterized by a focus on the understandings that result from a particular set of actions in a specific context of investigation. As such scientific inquiry involves an integration of a set of skills utilized within the laboratory and substantive knowledge addressing the same context. Specifically Gott and Duggan (2002) develop a set of concepts, termed concepts of evidence, that describe the types of understanding of action that they consider useful in characterizing scientific inquiry. This list consists of

• “Instruments: underlying relationships
• Instruments: calibration and error
• Reliability and validity of a single measurement
• The choice of an instrument for measuring datum
• Design: Variable structure
• Design: Validity, ‘fair tests’ and controls
• Design: Choosing values
• Design: Accuracy and precision
• Design: tables
• Reliability and validity of the design
• Data presentation
• Statistical treatment of measurements of data
• Patterns and relationships in data
• Reliability and validity of the data in the whole investigation” (Gott and Duggan, 2002, p. 117).

Roberts and Gott (2003), while careful not to associate the written test of concepts of evidence as an assessment of the performance of scientific inquiry, do
suggest that their list of concepts of evidence can be used in a written format and provide useful information on the knowledge-based aspects of scientific inquiry.

Solano-Flores and Shavelson (1997) and Shavelson et al. (1998) have developed an approach to the assessment of scientific inquiry, termed science performance assessment (SPA), which is based on the presentation of structured performance tasks to students in conjunction with a clearly formulated scoring system. As stated by Shavelson et al. (1998) “SPA should put a student in a laboratory, pose a problem, and watch as the student devises procedures for carrying out an investigation, analyzes data, draws inferences by linking data to prior knowledge (e.g., relevant theory), and comes to a conclusion and a problem solution” (p. 171). These researchers divide the structured tasks that they create for assessment purposes into four different types:

- **Comparative investigations** (scientific inquiries that “compare two or more objects on some attribute while controlling other variables” (p. 172));
- **Component-identification investigations** (scientific inquiries that require the determination of the parts that make up the whole);
- **Classification investigations** (scientific inquiries that require the development of a classification scheme utilizing the attributes of a set of objects);
- **Observation investigation** (scientific inquiries that involve the recording of a systematic observation over a period of time).

A core aspect of these researchers’ approach to assessment is the development of a reliable scoring system that addresses the central aspects of each inquiry type. For comparative investigation, the way the procedure is constructed and the accuracy of the problem are considered central. For a component-identification investigation confirming and disconfirming the presence of components is considered central. For a classification investigation usage of object attributes is considered central. For the observation investigation accuracy of results based on the observation and the development of explanatory models is central. The actual manifestation of science performance tasks consists of the construction of a specific task according to the guidelines of one of the research types, the performance of this task by the student, and the assessment of the task through the observation of the task in accordance with the guidelines of the scoring system.

In a later paper, Solano-Flores et al. (1999) widen their concept of science performance assessments by proposing a “shell” for the design of new performance assessments. These researchers define shells as “blueprints that provide directions for assessment developers to generate reliable, valid PAs [performance assessments] in a short time” (Solano-Flores et al., 1999, p. 294). The shell format builds upon the previous work conducted by this group of researchers and presented in Solano-Flores and Shavelson (1997) and Shavelson et al. (1998). Working within the scheme of the four inquiry types reviewed above, the shell divides the task into four student stages: planning, hands-on investigation, analysis, and interpretation and application. The shell also allows the manipulation of the task according to the type of equipment, the number of variables, the amount of conceptual information, the directions given, and the use of equipment. Through the manipulation of these
Assessing Scientific Inquiry

Factors the inquiry tasks can be conceptually simpler or more difficult. The manifestation of this approach is similar to that of the previous assessments conducted by these researchers. A task is presented to the student and this is assessed through assessment of student notebooks using the developed scoring system.

Using a more modest approach, Lunsford and Melear (2004) propose that the way to evaluate scientific inquiry is through a close consideration of the outcome of the scientific inquiry process. In their approach they avoid directly assessing the process of scientific inquiry and rather suggest that the final product of the inquiry can be used to infer what the student learnt and what was actually done in the project.

Within this approach a variety of different outcomes could be assessed. They suggest the use of portfolios (observation notes, data, etc.), scientific research papers, laboratory practicals and student demonstrations, student journals, and concept maps. For each of these products a specific set of assessment criteria have been proposed. The criteria are based on the genre aspects of the form of writing and components that are found within each of these written products. For example a research report is evaluated for the presence of a research question; a hypothesis; a literature review; a methods section; data presentation in tables, charts, and graphs; usage of references; references in paper; a bibliography; good grammar; and punctuality.

This list of approaches is not intended as a comprehensive meta-analysis of all the work that has been done on the issue of assessment and scientific inquiry. But it does capture the conceptual directions that have been proposed in different educational settings. Broadly the approaches presented above can be differentiated into two groups: those methods which assess scientific inquiry as an abstracted form of knowledge that could inform a range of studies and those which require students to perform a structured scientific task. The latter group can be further differentiated between the assessment of practical knowledge of instruments and measures and those that require different levels of problem solving. A review of all these approaches to assessment and the way they define scientific inquiry presents an array of knowledge types that collectively seem to cover the nature of scientific inquiry. These types of knowledge and abilities can be divided according to the following scheme:

1. **Substantive Knowledge**: Knowledge of scientific concepts, facts, and processes; knowledge of the stages and components of a scientific inquiry
2. **Procedural Knowledge**: Knowledge of procedural aspects of how to conduct a scientific inquiry; knowledge of the equipment and materials within a scientific inquiry
3. **Problem Solving and Integrative Abilities**: The ability to solve problems, pose solutions, conceptualize results, and reach conclusions

Based on this review, scientific inquiry would seem to be a multifaceted activity that requires several different levels and types of knowledge that need to be used in an integrative and problem-solving manner.

All the ways of assessing reviewed here in varying degrees rely on indirect, inauthentic measures of scientific task performance. This results in a large extent from the mass education school age framework of the educational settings that they
address. In-school scientific inquiry poses a series of problems such as large class sizes; the requirement to address national, state, and federal educational authorities; lack of equipment; difficulties in current scientific knowledge; classroom management issues; and the emphasis within schools on established knowledge that impede the use of authentic scientific inquiry organized around disciplinary significant research questions. For the authors of this book and the direction that is proposed here, authenticity is a serious issue.

Messick (1994), in a discussion of authentic assessment, points out that one of the problems in relation to validity of assessment measures is that of construct underrepresentation. As defined by Messick (1994) construct underrepresentation is a situation in which the authentic characteristics of tasks are not captured by the mode of assessment chosen. The multifaceted nature of scientific discourse, as well as the indirect, inauthentic way that it is assessed, poses the problem of construct underrepresentation in relation to the assessment of scientific inquiry. To have a valid assessment of scientific inquiry the different facets of authentic scientific inquiry need to be covered.

### 4.3 Authentic Assessment

A central aspect of authentic assessment is it situates the knowledge to be assessed within its social, physical, conceptual, and disciplinary contexts. As defined by Cumming and Maxwell (1999) authentic assessment considers real tasks in their context of performance and requires the person who is being assessed to apply expert problem-solving skills to the task on hand. Performance assessment focuses specifically on the completion of a specific task. What makes this approach authentic assessment is that the task requires the integration of disciplinary (contextually) specific knowledge in order to complete the task. Shavelson, Baxter, and Pine (1992) specify that performance testing requires students to perform concrete meaningful tasks that are assessed in relation to the types of reasoning and thinking that are applied to the task and not just the outcome. Stenmark (1991), in an early and rather open description, defines performance assessment as a procedure which involves “presenting a student with a task, project or investigation and then observing, interviewing, and looking at their products to assess what they actually know” (p. 13).

Herrington and Herrington (2006), summarizing the work of a series of researchers, offer the following set of characteristics that define the nature of authentic assessment:

1. Authentic assessment should involve an accurate recreation of the conditions under which the performance of the tasks would occur in its authentic setting (Reeves & Okey, 1996; Meyer, 1992; Wiggins, 1993).
2. Authentic assessment involves problem-solving skills and higher-order thinking (Reeves 2000; Newman & Wehlage, 1993)
3. Authentic assessment involves the production of new knowledge and not just the memorization and reproduction of established understandings (Newman & Archbald, 1992).


5. Authentic assessment is fully integrated with the activities that are the focus of the learning process (Reeves & Okey, 1996; Young, 1995).

As described above, authentic assessment implies a correspondence between the assessment task and the professional, real-world performance of the same task. From an assessment perspective, Gulikers (2006) provides an important criterion for evaluating the authenticity of an authentic assessment task when she states that “the authenticity of an assessment is defined by its resemblance to the real world” (p. 22). She defines this authentic, real-world task (termed the criterion situation) as the point of comparison for the assessment task. The degree of resemblance between these two manifestations of this task is the measure of the authenticity of the task. On this basis her definition of authentic assessment task is “an assessment that requires students to use the same competencies, or combinations of knowledge, skills and attitudes that they need to apply in the criterion situation in professional life” (p. 23). This definition of authentic assessment addresses a range of dimensions that function in relation to any real-world task. These include the nature of the task itself, the physical context of the task, the social setting and forms of professional interaction, the professional outcomes of this task, and the real-world criteria that are used within professional practice in relation to this task.

For the purposes of the approach to the assessment of scientific inquiry developed in this book, it is important to specify the ramifications of authentic assessment on the educational program. The first ramification deals with the concept of the construct validity of an assessment. Construct validity is defined as the degree to which an assessment task actually measures what it is supposed to measure. Since authentic assessment has a high degree of resemblance to the criterion situation and recreates the professional aspects of conducting the task under normal circumstances, authentic assessment tasks have very high levels of construct validity. Accordingly, the assessment of knowledge through an authentic assessment task should be a good predictor of the way the same task would be conducted within a professional setting. In other words, if one of the main aims of in-laboratory science instruction is to provide students with an experience of professional science and the instructor with some indicator of who has potential as a scientist, authentic assessment provides data on performance with an authentic laboratory task.

The second ramification is the inherent motivation and meaning that is found within authentic assessment tasks. Several researchers have specified that authentic assessment engages students and has a positive effect on student learning and motivation (Herrington & Herrington, 1998; McDowell, 1995; Sambell, McDowell, & Brown, 1997). As a result of the authentic contextualized nature of the performance of the task utilized for assessment, the procedure itself becomes part of the
instructional design and provides diagnostic information that is directly applicable to the student’s learning and development (Darling-Hammond & Falk, 1997). Authentic performance assessment has been recommended as particularly suitable for the development of scientific concepts and the enhancement of communicational skills required for problem solving in this context (Shymansky et al. 1997). This approach to assessment can legitimize the use of a wide variety of sources of information that provide evidence of both the process and outcomes of a student’s learning process. Performance assessment really allows teachers to focus on students’ higher-order thinking and reasoning. Finally and most importantly science students who were involved in performance assessment were found to have more metacognition of their reasoning processes and conceptualizations than students who underwent traditional testing procedures (Firestone Mayrowetz, & Fairman, 1998). In addition, authentic assessment which models professional activity is seen by students and instructors as a stage toward professional activity and not as an artificial measure of ability. As such, authentic assessment has the potential to be meaningful to the student on a personal level and provides the student with the option of really judging what the professional experience of science is like.

4.4 The Characteristics of Authentic Scientific Inquiry Assessment (ASIA)

As specified in the first chapter of this book, the context of interest for the current discussion of assessment is the in-laboratory scientific inquiry educational program. This educational setting is different from those that the previous approaches have addressed. However, from a conceptual perspective, the approach described here builds upon previous understandings of the ways to assess scientific inquiry. Specifically this approach builds upon the concepts of authentic assessment, the multifaceted nature of scientific inquiry, and the importance of carefully defining performance assessments.

The core characteristic of the approach described here is the importance it places upon authenticity in the educational program and the assessment of knowledge. For this reason the approach is best termed as authentic scientific inquiry assessment (ASIA).

As discussed in the previous section, authenticity in assessment is measured through the degree of comparison between the situation criterion (the real-world situation) and the assessment task. For an assessment task to be considered authentic it needs to have a high degree of equivalence with the tasks, the physical context, the social setting, forms of professional interaction, the professional outcomes of this task, and the real-world criteria that are used within professional practice in relation to this task. This form of assessment also requires students to use the same competencies or combinations of knowledge, skills, and attitudes that would be used in the professional setting.
The setting for assessment that the current discussion deals with is the in-laboratory, scientific inquiry, educational program directed by the presence of a research question that is significant within the wider scientific community. This specific educational setting allows high degrees of authenticity as it is situated in a laboratory and organized around real scientific research. The real issue for the development of an assessment program is to be able to accurately describe the nature of the educational, scientific activity and identify points of significance and value that constitute qualitative shifts in the understanding and practice of this scientific inquiry. In other words, the assessment procedures need to be able to accurately reflect the significant aspects of the scientific inquiry process that is at the heart of the educational program.

The answer that is at the center of the process of assessment development within the ASIA approach is the proposition that assessment development is preceded by empirical description of the scientific inquiry program. As seen in the approaches described in a previous section of this chapter, the definition of scientific inquiry was an early stage of most assessment development programs. The emphasis in ASIA is on the description of the knowledge, procedures, and representations that construct a specific scientific inquiry that is being used within the in-laboratory educational program. The assumption is that authentic scientific inquiry is a contextualized process that requires specialized knowledge. Accordingly, the starting point of assessment development is the empirical description of the specific aspects of this scientific inquiry.

The practicalities of an initial empirical description of a scientific inquiry that is relevant for assessment development pose a problem for most active scientist-educators. Accordingly the active assessment approach provides an analytical framework through which specific scientific inquiry programs can be analyzed in a manner that produces valuable distinctions. The next chapter provides a full description of this framework. From a conceptual assessment perspective, the approach used here involves a reversal of direction of the generative shell proposition for assessment development proposed by Solano-Flores et al. (1999). In the proposition of Solano-Flores et al. (1999) an analytic shell describing types of scientific inquiry tasks provided a framework for the production of new tasks following the same underpinning definitions. In the active assessment proposal, this analytic frame is used by scientist-educators to analyze the process of scientific inquiry that they are using in their laboratory. This analysis of scientific inquiry process produces a set of distinctions and categories of knowledge that can be used to define specific assessment tasks. The overall aim of this stage is to provide a theoretical framework for analyzing and characterizing the tasks that collectively define the scientific process that will be assessed. Since these assessment tasks are based on a process of empirical description conducted by the scientist-educator and drawn directly from the specific scientific inquiry utilized within the educational program, it is assumed that they will have a very high degree of authenticity.

An additional aspect of the ASIA proposal is the attempt for comprehensiveness in assessment. As seen in the previous section, scientific inquiry involves different types of knowledge and abilities. In order to capture the authentic aspects of the
specific scientific inquiry used within the in-laboratory educational program it is assumed that assessment will need to be multifaceted, involve several different elicitation formats and be conducted at different times. In this sense, ASIA is a process that both follows and is integrated within the actual experience of scientific inquiry with the in-laboratory scientific experience.

The characteristics of ASIA can be summarized as follows:

1. **Authenticity Through Empirical Description**: A high degree of equivalence is required between the situation criterion and the assessment tasks resulting from an initial stage of assessment development that consists of the empirical description of the specific scientific inquiry process that is at the core of the scientific agenda of the laboratory and the educational process.

2. **Multimethod Approach**: Since scientific inquiry is considered to be multifaceted, in order to capture the different aspects of scientific inquiry an assessment program will need to involve different components using different elicitation methods.

### 4.5 Active Assessment Development

The active assessment development procedure consists of the following stages:

1. **Empirical Description of Scientific Inquiry**: The first stage of assessment development consists of in-depth qualitative research that explores the nature of the scientific process being investigated. The analysis of the process is conducted using a developed framework for the analysis of the process of scientific inquiry and the types of knowledge and abilities required in this inquiry process. The aim of this stage is to characterize the tasks that collectively define the scientific process that will be assessed.

2. **Definition of Educational Aims**: Following the empirical description of the scientific process that is utilized within the educational, scientific inquiry program, the educational aims of the program that the scientist-educator wants to assess need to be defined. A special aspect of the active assessment approach is the idea that the educational aims of program and the definition of significant scientific educational development are directly tied to the scientist-educators’ understanding of the scientific inquiry process at the heart of the educational program. In other words, the definition of what is important to evaluate within the educational program results from the understanding and analysis of what is important within the scientific inquiry process itself. Specifically, in relation to active assessment, the scientist-educator will have to decide which of the various representational milestones (and associated cognitive and physical knowledge) is to be assessed. It is assumed that the whole process of assessment will address summative, formative, and diagnostic aims and will address representational, cognitive, physical, and presentational knowledge.
3. *Assessment Tool Development*: Once the scientific process has been characterized and the significant aspects of the process decided upon, the process of tool development can start. A core principle of this stage is the idea that the assessments need to be as authentic as possible and are directly tied to the analysis of the scientific inquiry process conducted as a first stage of active assessment. The development of specific assessment tools is the attempt to model directly the aspects of conducting the scientific inquiry itself. There are different ways of doing this, and tools of this kind could come in different elicitation formats. However, at every point it is important to understand that the tools should be directly tied and modeled on the analysis of the specific inquiry process that is being used in the educational program. The set of tools developed should comprehensively cover all the different types of knowledge used by professional researchers in conducting this type of scientific inquiry. In order for the assessment tools to be authentic they will need to integrate physical, representational, cognitive, and presentational knowledge.

4. *Scoring Rubric Development*: Having created a set of tools through which knowledge will be assessed in the educational program, a set of scoring rubrics need to be developed. As with the development of the tools themselves, a core aspect of the active assessment approach is that the scoring of the tasks is as authentic as possible. In other words, the criteria for success are drawn from the context of conducting the same tasks within the professional setting. The scoring rubrics should be defined according to the professional knowledge of the scientist-educator. As with all scoring rubrics, criteria for successful and problematic outcomes on these tasks need to be defined. The scoring rubrics provide the educational team and the student-researcher with both formative and diagnostic information that can enhance the educational program.

5. *Assessment Piloting*: As with any other assessment process, it is important to pilot the developed assessment tools. The process of piloting consists of using the developed tools with a smaller subset of the population and the constant reflective evaluation of the ability of the tools to assess the defined educational aims of the program. The stages of tool development, rubric development, and piloting are closely interrelated. A core assumption of active assessment is that piloting stage will evolve and change the tools that are used and the way they are scored. The outcome of this whole process should be an assessment strategy consisting of a series of piloted tools, which is authentic to the specific scientific process that is being taught, which is meaningful to the student and instructors, and which can be scored in a reliable, systematic manner.

### 4.6 Chapter Summary

The aim of this chapter was to review proposals concerning ways of assessing scientific inquiry and to propose a specific process for the assessment of scientific inquiry termed authentic scientific inquiry assessment (ASIA). The following ideas and concepts were defined:
1. A range of assessment procedures have been developed to assess scientific inquiry. These procedures differ in relation to the way they define scientific inquiry and the elicitation methods that they employ. These procedures rely on indirect, inauthentic measures of scientific inquiry.

2. A comprehensive analysis of the different definitions of scientific inquiry for assessment purposes defines scientific inquiry as multifaceted and involving substantive knowledge of scientific facts, concepts, and process, substantive knowledge of the stages and components of scientific inquiry, procedural knowledge of how to conduct a scientific inquiry, procedural knowledge of equipment and materials, and problem-solving abilities.

3. Authentic assessment creates assessment procedures that are closely modeled on real-world tasks and require the same conceptual, physical, and social abilities as the real-world task.


5. ASIA is an approach that is characterized through its definition of assessment tasks by a process of empirical description that enhances the authenticity of the assessment procedures and its emphasis on comprehensive multifaceted assessment of scientific inquiry.

6. The process of active assessment development consists of five stages: empirical description, definition of aims, tool development, rubric development, and assessment piloting.
Chapter 5
An Analytical Framework for the Development of Scientific Inquiry Assessment

5.1 Introduction
The aim of this chapter is to provide scientist-educators with a set of conceptual tools that can be used to develop an assessment program for in-lab, educational, scientific inquiry programs. This chapter operationalizes the concept of authentic scientific inquiry assessment (ASIA) within the active assessment approach. As argued in previous chapters, a core assumption of the active assessment approach is the idea that the prime researcher within a professional research laboratory is directly involved in the design of both the educational scientific inquiry program and the way this program is assessed. Accordingly, this chapter is designed to provide the scientist-educator with a terminology and conceptual methodology to fulfill this function.

5.2 Stages of Active Assessment Development
As described in the previous chapter, the process of developing an assessment program for scientific inquiry involves five stages: empirical description, definition of aims, tool development, rubric development, and assessment piloting. Figure 5.1 schematically represents the active assessment development process.

As discussed in the last chapter and schematically represented in Fig. 5.1, the core aspect of the active assessment approach to the development of authentic scientific inquiry assessment is the significance assigned to the empirical description of the scientific inquiry process. Active assessment aims to develop an overall assessment approach that is as authentic and as meaningful as possible. Accordingly the central task of assessment development is the explicit description of the scientific inquiry process that can form the basis of the assessment tasks. The authors of this book recognize that one of the problems that this poses to the scientist-educator is the actual conceptual organization of this knowledge in a format that can generate assessment tools. The sections that follow are designed to provide this conceptual framework through which the aims of active assessment can be achieved.
5.2.1 Stage 1: Empirical Description of Scientific Inquiry

The first stage of the active assessment process consists of the empirical description of the scientific inquiry process which constitutes the educational program. In some ways, the aim of this section, in helping professional scientists to define their scientific inquiry processes, is paradoxical. As professional scientists you are well aware of your own scientific inquiry processes and for your own professional activities you are not in need of any help from the authors of this book. However, in developing an assessment program, the conceptual framework developed here may be useful for its ability to create distinctions in the types of knowledge required and the ways these are integrated and organized across the scientific inquiry process. As will be described in the next section of this book, our experience within our microbiology laboratory has shown that this way of describing the scientific inquiry process is a useful way to generate a set of assessment tools and an overall strategy toward the assessment of our program.

The conceptual framework is constructed from the model of scientific inquiry proposed in Chap. 2 of this book and consists of a series of directed questions which should allow a clear description of a specific scientific inquiry process to emerge. As described in Chap. 2, the description of a specific scientific inquiry process consists of the analysis of two axes: Axis 1 consists of an analysis of the stages of development of the scientific inquiry process; Axis 2 consists of the types of knowledge required in order to complete the scientific inquiry process. As seen in Chap. 2 this is schematically represented in Fig. 5.2.

As seen in Fig. 5.2, the core assumption of this model is that professional scientific inquiry is an extended process and consists of a series of stages with concrete outcomes. Each of these stages and outcomes is seen as a form of “milestone” in the process of scientific inquiry and when accumulated these stages of the scientific inquiry process lead to potentially significant findings. Within the scientific inquiry process each of these stages or outcomes is seen through the presence of a particular form of visual representation. This representation might be as diverse as a
fluctuation on a computer readout or a plaque community on a dish. But there is some form of visual representation that marks the completion of a stage of the inquiry. Accordingly, these representations can be used to divide the scientific inquiry process into stages. For each of these stages specific physical manipulations will need to be done. In other words, the outcomes are dependent on some form of physical manipulation. For these manipulations and representational outcomes to be meaningful, the researcher needs to understand the relevant substantive knowledge and associated procedural understandings. Taken together, the cognitive, physical, and representational knowledge types organized along the developmental line of the scientific inquiry process should produce a description of the specific inquiry process used in the educational program.

The conceptual framework presented here consists of a series of questions asked and answered by the prime researcher and designer of the educational, scientific inquiry program. These questions are to be thought of as a form of reflective exercise in which the scientist-educator explores her/his own knowledge of the scientific inquiry process. It is assumed by the authors of this book that most scientists will be very well acquainted with their own research agendas and associated knowledge types. What this reflective exercise provides is a way of explicating, categorizing, and organizing this implicit knowledge base.

The first set of questions, designed according to the two axes of scientific inquiry, is planned to explicate the process itself and consists of the following:

1. **Axis 1: Representational Knowledge**: What types of representation are produced within the scientific inquiry process?
2. **Axis 2: Inquiry Process Development**: Is there a specific expected order for the production of these representations? What is that order (or potential orders)?
3. **Axis 1: Physical Knowledge**: What physical procedures are required to produce these representations?
4. Axis 1: Cognitive Knowledge: What scientific content knowledge is required to understand and interpret these representations? What informed decisions (or calculations) need to be made in order to continue this process of scientific inquiry? What does one need to know in order to understand the different representation produced within the scientific inquiry process?

5. Axis 1: Presentational Knowledge: In what ways and in what settings are the results of this scientific inquiry presented to the community of scientists?

The analytical process starts with a consideration of the specific representations found in the scientific inquiry process. The first question of this stage provides the basis for the division of the scientific inquiry process into stages. By focusing on concrete outcomes and their representational form a basis is created for the design of later assessment measures.

Having defined the specific representations found in this scientific process the second question deals with the way these representations organize the development of the scientific inquiry process. By focusing on the order of the representation the overall scientific inquiry process is divided into distinct stages according to the production of a specific outcome that is represented in a particular way. Through questions 1 and 2, the specific representations that are expected and the order in which they are expected create a timeline of the specific scientific inquiry that is being taught.

For some professional scientific inquiry programs, it is possible that the last stages of the inquiry cannot be determined at the beginning of the process, in that the scientific inquiry process is open ended (with every finding leading to yet another set of research questions). Accordingly, for some in-lab, educational, scientific inquiry programs it is assumed that only the first two-thirds of the program can be mapped out according to known representations.

Once the first two questions have been answered and a timeline of representations has been created, questions 3 and 4 address the physical and cognitive knowledge required to produce and understand each of these representations. The physical procedures that need to be assessed may include the working of equipment, ascetic practice, safety behaviors, and a wide range of laboratory work. For each representation, the physical actions which lead to this outcome need to be noted. In order to produce the specific representation, relevant scientific knowledge needs to be known. The prime researcher should consider what scientific knowledge needs to be known in order to understand what is happening that produces the required result from a scientific perspective. This will include factual knowledge of the processes and factors that influence any given outcome. In addition, the prime researcher should consider the problem-solving aspects associated with a particular stage of the scientific inquiry. What decisions need to be made? What are the range of options for interpreting the results found? What are the ways in which a particular representation can be understood? For most scientists the main problem with questions 3 and 4 is recognizing the depth of their contextualized scientific knowledge. Novice researchers will know far less than the expert and may often surprise researchers with questions that are truly basic. When considering the
knowledge required in order to complete and understand a stage of a scientific inquiry, it is worth taking a very conservative approach to the amount of knowledge that students may have and to consider very basic as well as more sophisticated knowledge that contributes to the understanding of the process. The definition of physical and cognitive knowledge needs to be given for each of the representation along the timeline of the scientific inquiry process. The outcome of questions 1–4 should be a detailed description of a scientific inquiry process that includes physical, cognitive, and representational knowledge and is organized along a timeline of representations. The fifth question addresses provisional outcomes of the inquiry process and explores the ways in which scientific knowledge is commonly presented within the specific discipline.

5.2.2 Stage 2 – Definition of Aims

Having developed an explicit description of the scientific inquiry process, the second stage of active assessment development consists of deciding on the aims of assessment and the definition of significant scientific knowledge. During stage two, types of decision need to be made. The first decision addresses the relationship between the assessment and the objectives of the educational program. The second set of decisions addresses the definition of knowledge that is to be considered significant within the specific scientific inquiry process used in the educational program. Together these two sets of decisions define the purpose and content of the assessment program.

The first decision relates to the characteristics of the educational program itself. The prime scientist-educator needs to consider what the educational program is really trying to achieve from an educational perspective. The aims of the assessment program are tied directly to the overall objectives of the educational program. As such these aims need to be explicitly expressed so that the role of assessment in providing evidence that assess these aims can be defined. This stage of active assessment development is directed by the following two questions:

1. What are the aims of the educational program?
2. What are the aims and purposes of the assessment program?

In most scientific inquiry programs, the aims of assessment are summative, formative, and diagnostic. The aims of assessment are to provide evidence on the outcomes of learning (summative) and to improve instruction and facilitate student learning (formative and diagnostic). In science education, there may be specific aims such as the development of scientific literacy or acquaintance with a specific scientific discipline. The aims of the assessment program need to be explicitly and clearly written out. From a methodological perspective, these aims function as a set of research questions that direct the process of data collection in the form of assessment.
The second set of decisions deals with the content of scientific knowledge that is considered important for the specific scientific inquiry project that is utilized within the educational program. This stage is directly informed by the empirical description developed in the first stage of active assessment development and consists of deciding on the hierarchy of knowledge within the scientific inquiry process. A basic assumption of this stage is that the assessment process will address significant knowledge that is crucial for understanding this specific scientific inquiry. Accordingly, decisions need to be made as to what consists of significant knowledge within this specific scientific inquiry process. This aspect of active assessment is directed by the following questions:

1. Of the representations defined in this scientific process (see question 1 above), which of these representations could be considered a representational “milestone” in the development of this scientific process? (A limited set of crucial, milestone representations in the development of a scientific inquiry need to be chosen.)
2. What physical laboratory knowledge, cognitive scientific knowledge, cognitive interpretive and decision-making knowledge, and knowledge of visual representation are required to create and understand the milestone representations that have been chosen (see previous question)? (The required knowledge in each case needs to be explicated as fully as possible.)

The outcome of this analytical process is the decision that a subset of stages within the scientific inquiry process is of greater significance from a knowledge perspective than others in this specific scientific inquiry process. Each of the stages should be considered as a significant moment within the timeline of the scientific inquiry that can be used as a measure of the student-researchers’ development of understanding of the scientific inquiry process.

### 5.2.3 Stage 3: Tool Development

The aim of this active assessment stage is the actual development of tools that can be used with student-researchers in order to assess knowledge development. As discussed in previous chapters, scientific inquiry is assumed to be multifaceted and as such a multimethod approach will be required to assess its different aspects. In addition, a core assumption of active assessment is the value placed on authenticity. There are several considerations that need to be addressed during the development of assessment tools:

- Comprehensively address representational, cognitive, physical, and presentational knowledge relevant to this specific scientific inquiry process
- Address significant knowledge within the scientific inquiry process
- Address the development of knowledge along the timeline of the scientific inquiry process
- Be as authentic as possible in designing the assessment tools
Make sure that the assessment tools can actually provide evidence that is useful in addressing the aims and purposes of the assessment program.

The development of assessment tools is a creative, problem-solving activity. The aims and scientific content of the assessment tools have been defined in previous stages. But these definitions need to be operationalized at this stage in the active assessment development process. To a certain extent, the process described here is simpler than those used by many other assessment development programs. As a rule of thumb, the direction taken in the active assessment approach is to use authentic aspects of the scientific inquiry process whenever possible. In other words, a large part of the tool development process is directed by the description of the scientific inquiry process (stage 1) and the decisions as to which aspects of this process are considered significant (stage 2). The development of assessment tools is relatively simple because it draws directly on the contextualized knowledge required for the completion of the scientific inquiry process itself and is based on the in-depth understanding of the scientist-educator of her/his own scientific inquiry process. In line with the concept of authentic performance-based assessment the actual assessment tools are actually tasks from within the scientific inquiry process that have been modeled. It is assumed that each of the significant milestones of the scientific process and the associated knowledge sources will be assessed. For some programs, this may mean actually observing the students’ way of conducting the task; for others it may mean creating more indirect ways of eliciting the same knowledge. In this sense, assessment tasks can be modeled on the authentic tasks used in the scientific inquiry process. These may include completion of a laboratory process, the explanation of laboratory result, the writing and reading of a laboratory notebook, and the presentation of a poster. For active assessment, the idea is that actual authentic aspects of professional scientific inquiry be used as moments of student knowledge assessment.

When using authentic assessment tasks, it is important to make sure that all types of knowledge are assessed. In particular it is important that physical, cognitive, and representational knowledge are integrated in the evaluation of scientific inquiry. This will probably mean that the assessment tools will be of different types that use different methods of data elicitation. This multimethod approach should ensure that collectively the different types of knowledge comprehensively cover the different types of knowledge that are being assessed. This may manifest itself in an assessment program that consists of some performance of specific tasks, some indirect elicitation of cognitive knowledge, and some summarizing presentation task at the end of the scientific inquiry process. The tools of assessment need to be carefully designed and explicitly described and written.

5.2.4 Rubric Development

The aim of this stage is the development of scoring rubrics that can be used to rate the outcomes of student performance on the different assessment tools. In
Table 5.1  Summary of main aspects of active assessment development

<table>
<thead>
<tr>
<th>Stage</th>
<th>Aim</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Empirical</td>
<td>Description The aim of this stage is to develop an explicit</td>
<td><strong>Axis 1: Representational Knowledge:</strong> What types of representations are produced within the scientific inquiry process?</td>
</tr>
<tr>
<td></td>
<td>description of the scientific inquiry process that is used within the educational program</td>
<td><strong>Axis 2: Inquiry Process Development:</strong> Is there a specific expected order for the production of these representations? What is that order (or potential orders)?</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Axis 1: Physical Knowledge:</strong> What physical procedures are required to produce these representations?</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Axis 1: Cognitive Knowledge:</strong> What scientific content knowledge is required to understand and interpret these representations? What informed decisions (or calculations) need to be made in order to continue this process of scientific inquiry? What does one need to know in order to understand the different representation produced within the scientific inquiry process?</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Axis 1: Presentational Knowledge:</strong> In what ways and in what settings are the results of this scientific inquiry presented to the community of scientists?</td>
</tr>
<tr>
<td>2. Definition of Aims</td>
<td>The aim of this stage is to explicitly define the purpose of the assessment program and set priorities in relation to what is significant scientific knowledge in this specific scientific inquiry process</td>
<td>What are the aims of the educational program?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are the aims and purposes of the assessment program?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Of the representations defined in this scientific process which of these representations could be considered a representational “milestone” in the development of this scientific process? (A limited set of crucial, milestone representations in the development of a scientific inquiry need to be chosen)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What physical laboratory knowledge, cognitive scientific knowledge, cognitive interpretive and decision-making knowledge, and knowledge of visual representation are required in order to create and understand the milestone representations that have been chosen (see previous question)? (The required knowledge in each case needs to be explicated as fully as possible)</td>
</tr>
</tbody>
</table>

(continued)
### Table 5.1 (continued)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Aim</th>
<th>Questions</th>
</tr>
</thead>
</table>
| 3. Tool Development | The aim of this active assessment stage is the actual development of tools that can be used with student-researchers in order to assess knowledge development | The direction taken in the active assessment approach is to use authentic aspects of the scientific inquiry process whenever possible Tool development process is directed by the description of the scientific inquiry process (stage 1) and the decisions as to which aspects of this process are considered significant (stage 2)  
  **Considerations:**  
  (a) Comprehensively address representational, cognitive, physical, and presentational knowledge relevant to this specific scientific inquiry process  
  (b) Address significant knowledge within the scientific inquiry process  
  (c) Address the development of knowledge along the timeline of the scientific inquiry process  
  (d) Be as authentic as possible in designing the assessment tools  
  (e) Make sure that the assessment tools can actually provide evidence that is useful in addressing the aims and purposes of the assessment program |
| 4. Rubric Development | The aim of this stage is the development of scoring rubrics that can be used to rate the outcomes of student performance on the different assessment tools | Scoring rubrics are based directly on the professional knowledge in conducting these tasks within the authentic setting |
| 5. Assessment Piloting | The aim of this stage is to improve and validate the assessment procedures | To what degree do the assessment tools and the scoring rubrics provide evidence that can be used to assess the aims and purposes of the educational program? |
accordance with the principles of authentic assessment it is assumed that the scoring rubrics are based directly on the professional knowledge in conducting these tasks within the authentic setting. In other words, the criteria for success are identical to the criteria used by professional scientists in conducting the same task. The scoring rubrics are a crucial part of the assessment program because the data elicited from the students is understood in relation to the standards of responses defined in the scoring rubric. Scoring rubrics usually divide responses into exemplary, acceptable, and unacceptable levels of response. The scoring rubrics need to be explicitly defined and written.

### 5.2.5 Assessment Piloting

The aim of this stage is to improve and validate the assessment procedures. All assessment tools and scoring rubrics need to be piloted. At the heart of the piloting program is the question

1. To what degree do the assessment tools and the scoring rubrics provide evidence that can be used to assess the aims and purposes of the educational program?

The process of piloting assessment tools requires the actual completion of the assessment tools by student-researchers and the scoring of these responses by different members of the educational team. It is a basic assumption of any piloting procedure that both the tasks and the scoring rubrics will be modified through the piloting process. In addition to the piloting of the materials with student-researchers, it is often useful to have professional scientists complete the assessment tasks as well and to look at the scoring rubrics. These professionals will ensure the authenticity of the materials developed and the relevance of these to the specific scientific inquiry process.

### 5.3 Chapter Summary

The aim of this chapter was to provide a conceptual framework for active assessment development. Table 5.1 summarizes the main aspects of this framework.
Chapter 6
The PHIRE Program

6.1 Introduction

The first section of this book (Chaps. 1–5) presents the theoretical basis for assessment of scientific inquiry and a practical framework for active assessment. This second section describes a case study and how the specific assessment strategies and tools were constructed and implemented. The first chapter of this second section presents a detailed description of the PHIRE program and establishes the context for authentic scientific inquiry and active assessment. The two subsequent chapters discuss the specific assessment strategies and the tools developed for active assessment within this program. The final chapter offers reflections from the perspective of the scientist-educator on the establishment of integrated research-education programs and the role of active assessment within this context. This case study is not presented with the expectation that it will be replicated elsewhere, but that it will provide a practical framework for establishing educational programs with fully integrated active assessment within the wide berth of sub-disciplines within the life sciences.

In this chapter we will provide a description of the PHIRE program and its development at the University of Pittsburgh. The acronym PHIRE is an abbreviation of the Phage Hunters Integrating Research and Education program and emerged several years after phage discovery activities begun in 2002 (further insights into the origins of the PHIRE program are provided in Chap. 9). The program name and acronym are rather informative. First, it places an immediate emphasis on the integration of scientific and educational missions, reflecting a core goal and motive for the program. Second, it introduces the concept of phage hunting, and while the general reader might know rather little about phages or the viruses of bacteria, it connotes the idea of searching and discovery. A key program theme is the idea of igniting enthusiasm in students through their involvement as research scientists, lighting a fire that will promote continued involvement in scientific research. It should also be noted that some of those involved in phage research often have a light-witted predilection to substitute “f” with “ph” at every opportunity!

A brief overview of the PHIRE program provides a context for a more detailed description. In this program, high-school and undergraduate students have the
opportunity to engage in a research program focusing on their discovery of new bacterial viruses (bacteriophages, or phages), followed by purification and characterization of the new viruses that they discovered, with emphasis on determining the genomic DNA sequence. The new phage genome sequence is then annotated to show how the genetic information is organized and compared with other genomes to elucidate its evolutionary history. There are a variety of variations on this theme and these will be discussed later.

To fully understand the PHIRE program and why we believe it has some special attributes facilitating an integration of our research and scientific missions, we will first present some helpful background information about bacteriophages and the size and diversity of the phage population. These scientific concepts are central to the PHIRE program, and being familiar with them is important for fully comprehending the advantages that the PHIRE program has to offer.

### 6.2 What Are Phages?

Bacteriophages are viruses that infect bacterial hosts. They are commonly referred to as phages and share all the typical properties associated with viruses. They are not capable of duplication or growth by themselves and can only multiply once they have infected bacterial host cells. They are small and their morphological details can only be elucidated by examination in the electron microscope (one example is shown in Fig. 6.1). Phages come in many different shapes and sizes, but most are composed of two readily identifiable structures, the head (or capsid) and a tail. The tail can of be various types, the three most common being a short stubby tail, a long flexible non-contractile tail, and a contractile tail; the tail type is used for taxonomic grouping, with these types being known as the Podoviridae, Siphoviridae, and Myoviridae, respectively. To give a sense of scale, the diameter of the head is usually 50–100 nM, and the length of the tail (in the Siphoviridae and Myoviridae) is typically in the range 70–300 nM.

Chemically, bacteriophages are composed of two major macromolecules, proteins and nucleic acids, although some have associated lipids and carbohydrates. The visible structures are made of protein, but there are usually many different specific types of proteins, sometimes 20 or more. The quantity of each of the individual proteins is broad, ranging from fewer than half-a-dozen to several hundreds of others, all within a single viral particle. The genetic information is carried as nucleic acid, with the majority of viruses carrying double-stranded DNA (dsDNA), although there are examples of those containing single-stranded DNA (ssDNA) or RNA. The DNA genome is carried inside the head of the phage, typically in a linear form, and remains there until the process of infection occurs. When the tip of the tail associates with a specific receptor on the surface of the host bacterial cell, the genome exits the head, traverses through the tail, and is injected across the cell wall into the cytoplasm of the cell; the viral protein shell remains attached to the outside of the cell and plays no further role.
6.2 What Are Phages?

Although phages are too small to see directly with a light microscope, they can be readily recognized and analyzed microbiologically. A common technique is called the “plaque assay,” in which a phage sample and bacterial cells are spread across an agar surface in a Petri dish and incubated so that bacterial growth generates an opaque lawn across the surface; with most bacterial types this takes 12–24 h. Whenever a phage particle in the initial sample infects a host cell, it can reproduce to generate more copies of itself, typically 50–200, which are then released by disintegration (lysis) of the infected cell. These progeny – which are identical to the initial particle – can then repeat this process and continue to do so as long as the bacterial cells are growing. What is generated is a plaque, a small circular area in which the bacteria have been killed and the phage multiplied (Fig. 6.2). At the point when the bacteria cease to grow (12–24 h), there are typically $10^6$–$10^7$ phage particles in a single plaque; all derivatives of the single particle that initiated that infection. This reproductive process requires the rapid and efficient production of many new copies of the phage genome and assembly of virion structures into which the DNA is packaged. The DNA replication process is accurate but not perfect, and thus all the progeny derived from a single particle will be extremely similar but not necessarily identical, with a small subset containing mutations making them just very slightly different to each other.
The plaque assay enables the simple recognition of a single viral particle in a sample, since one particle is sufficient for a plaque. Moreover, if there are many particles in a suspension, then through the process of purification by plating serial dilutions, it is possible to accurately separate and enumerate the phage particles in any sample. The number of particles is often referred to as its “titer,” and when the titer is determined in this way it is usually expressed as the number of plaque-forming units (pfu) per volume of sample. For example, if 0.1 ml of a $10^{-6}$ dilution (which can be achieved by three serial 1 in 100 dilutions of the initial stock) provides 13 plaques on a particular host, then the titer of the stock can be described as $1.3 \times 10^8$ pfu/ml.

Preparation of bacteriophages stocks can be readily accomplished by mixing bacterial host cells and phage particles and allowing multiple rounds of phage growth and cell lysis to occur. Eventually all or most of the cells will be lysed, and after filtration the resulting sterile “lysate” can be stored and used for future experiments. There infections can be performed on solid agar medium in a Petri dish to generate a “plate lysate” or in liquid medium to generate a “liquid lysate.” While either method can be used, some phages – including many of the mycobacteriophages – appear to grow better on solid medium than in liquid, thus generating higher phage titers. Either approach can generate stocks with very high titers ($10^9$–$10^{11}$ pfu/ml), although there are many parameters that influence this, especially the number of phage particles that are initially added to the cells. Optimal phage production often depends on empirically determining this for any given phage.

### 6.3 Bacteriophages Are Ubiquitous and Host Specific

Bacteriophages were first discovered separately by Felix D’Herelle and Frederick Twort in the early 20th century employing the plaque assay as described above. It became clear that phages could be identified in many types of samples from many different environments, but that each phage type is specific to the bacteria they can infect. The collection of bacteria that any specific phage isolate can infect is referred
to as its “host range,” which can only be determined empirically by asking what bacterial hosts it is capable of forming plaques on. Some phage types have rather broad host ranges and may for example infect many different bacterial species within a given genus. Other phages have extremely narrow host ranges and may infect only a specific sub-type of one bacterial species. Phages have been isolated for the majority of bacterial species for which they have been specifically sought, and it is generally accepted that phages probably exist for the vast majority of bacteria in the biosphere. It is noteworthy that only a rather small portion of bacteria in the environment can actually be cultured in the laboratory, and it seems probable that all of the unculturable bacteria are also hosts for bacteriophages.

6.4  **Bacteriophages: The Dark Matter of the Biological Universe**

In the early 1990s, methods were developed enabling the simple determination of the number of nucleic acid containing items in an environmental sample using only a fluorescence microscope (Hennes & Suttle, 1995). When certain dyes are added (e.g., SYBR Green), bacterial cells and viruses absorb the dye, binding to the DNA and causing them to fluoresce. When seawater samples were examined (which are easy to collect), it was found that they typically contain $10^6–10^7$ virus particles per milliliter, regardless of whether the samples were coastal, oceanic, surface, or deep (Wommack & Colwell, 2000); moreover, the ratio of viruses to bacteria (which are easily distinguishable as being larger than the viruses) was typically about 10:1. Similar numbers and ratios are estimated from terrestrial samples, and with a simple multiplication by the size of the oceans plus some factor for terrestrial areas, it can be calculated that there are about $10^{31}$ viral particles in the biosphere (Hendrix, 2002), of which the large majority are bacteriophages (i.e., they infect bacterial hosts). Since an independent calculation estimates that there are about $10^{30}$ bacterial cells in the biosphere (Whitman et al., 1998), this seems not unreasonable. However, the size of the phage population is truly stunning. The $10^{31}$ totality suggests that phages represent the numerical majority of all biological entities, i.e., there are more phage particles than all other forms of life taken together. If we assumed that each particle was on average 100 nm in length, then placing the entire population into a linear arrangement would generate an array that extends 200 million light years!

Recognizing the remarkable size of the phage population shows us how little we understand it. The total number of bacteriophages that have been described in the literature is perhaps a few thousand, and genomic analysis has been undertaken for just over 500 at the time of writing (September, 2008). We will provide further details into the comparative genomics of the phages in Sect. 6.7.8, but one simple conclusion is that they are genetically highly diverse; so these are not apparently $10^{31}$ particles that are minor variants of each other. In fact, the abundance of novel genes present in phage genomes suggests that the phage population represents the
largest unexplored reservoir of genetic information in the natural world (Hatfull et al., 2006).

6.5 Goals of the PHIRE Program

Realizing the size and diversity of the phage population provides a foundation for understanding the goals of the PHIRE program. The program emerged as a means of introducing students to the scientific process, emphasizing the involvement of students that have little scientific training but are curious about science and the natural world in which we live. Involvement of students in a research project within an established research environment provides a forum for an education into both the process and content of science. We therefore wanted to integrate our research and educational missions in a context that broadly extends opportunities to a diverse student body – both high school and undergraduate – without bias except for their curiosity and desire to learn about science (Hanauer, 2007). The specific research projects needed to address authentic scientific questions could accommodate a significant number of students and most importantly be intertwined with the excitement of scientific discovery. The diversity of the phage population and the ability to isolate novel phages relatively easily thus offered an attractive framework for engaging students in scientific research.

6.6 The Power of Discovery

The number and diversity of the phage population provides a core underpinning for the PHIRE program. While we would not make the claim that phage hunting is uniquely suited for scientific discovery by beginner scientists, it does appear to be particularly well suited to it. The genetic diversity of phages is sufficiently high that new isolates recovered from the environment are highly likely to be novel and, in many instances, quite different genetically from any viruses previously discovered. Furthermore, the isolation of a new “plaque-forming unit” is fairly simple from a technical perspective and doesn’t require specific training or special skills. It thus offers a simple entry point for anyone to begin the process of discovery, and students learn rather quickly that they can contribute to the scientific process. At the same time, it soon becomes apparent that the isolation and purification of a plaque-forming unit does not in itself provide any information about its novelty. Even if it is purified and examined by electron microscopy – an exciting moment of realization as to what the virus looks like – its genetic relationship to other phages is not known. It is only when the DNA has been extracted and the genome sequence determined that its relationship to known viruses is understood.

It is clear then that the process of phage discovery and genomic analysis is thoroughly imbued with scientific discovery. A “new” virus can be isolated and studied,
but its novelty is understood only by pursing the process of characterization. The
discovery – and the accompanying sense of project ownership – provides the impe-
tus and drive to continue and to open the door to further discoveries unearthed by
genomic characterization. Furthermore, the new genetic information derived is of
broad importance to the larger scientific community and thus warrants publication
in some form (Hatfull et al., 2006). It is an authentic piece of science.

6.7 PHIRE: A Ten-Step Program

Having established that the opportunities for scientific discovery by phage hunting
and genomic analysis are suitable for integrating the research and educational mis-
sions, we will describe the component steps of a typical project within the PHIRE
program. It will become apparent that there is a prescribed formula or shape to the
projects, each of which differs primarily by the individual phages that students iso-
late. We discuss in further detail in Sect. 6.11 the relative merits of such parallel
projects, but it provides a common platform in which we can identify ten individ-
ual steps, most of which follow each other in a logical and chronological order. We
will assume that the reader may have only a rudimentary knowledge of molecular
biology. A visualization of these steps is provided in Fig. 6.3.

6.7.1 Step 1: Phage Isolation

As noted above, a student joining the PHIRE program begins by isolating a bac-
teriophage from the environment. Step 1 involves the gathering of environmental
samples – we typically encourage collection from soil or compost although students
can be quite creative – in a 15 ml plastic tube. Back in the laboratory, samples are
extracted with a simple buffer (phage buffer) and filtered to remove contaminating
bacteria. A portion of this sample is then mixed with bacterial cells and thin agar
solution (top agar), and poured onto an agar base in a Petri dish. Once the top agar layer has solidified, the plate is incubated for 16–24 h and examined for the appearance of plaques on the bacterial lawn (see Fig. 6.2).

Most of our experience in phage hunting has been with using *Mycobacterium smegmatis* as a bacterial host although in principle any bacterial host can be used; we refer to phages that infect mycobacteria as mycobacteriophages. The proportion of samples containing identifiable plaques is relatively low (<10%) so that some students will be successful in their first set of samples, while others may have to continue to gather more samples before being successful. When plaques are observed, the numbers vary – often as few as just one but can be as many as hundreds. In the PHIRE program, close to 100% of students have succeeded in identifying at least one plaque. It should be noted that the samples probably contain large number of phage particles (>10⁶) but most of which infect bacterial species other than *M. smegmatis*. We have not systematically monitored what bacteria are present in the samples that students collect, although this has rich potential for future program development.

### 6.7.2 Step 2: Phage Purification

Step 2 involves the purification of the phage sample to ensure that it is a homogenous sample. This is important since without a clonal population further characterization becomes overly complicated and since you then can’t be sure of what you are really working with. The purification can be accomplished by “picking” the plaque by carefully stabbing it with a plastic pipette tip, transferring it to an aliquot of phage buffer, and re-plating dilutions to generate single plaques. With each plaque containing 10⁶–10⁷ particles, this can be fairly easily accomplished, provided that several dilutions are made to ensure that at least one gives a reasonable number of nicely isolated plaques. This is then repeated once or twice more to have confidence that a homogenous phage sample is achieved.

### 6.7.3 Step 3: Phage Amplification

Growth and preparation of a large phage stock is typically accomplished by first establishing the number of particles that will produce a maximum bacteriophage yield, followed by phage growth on solid medium using 30 large agar plates. After incubation the phage particles are harvested and concentrated by precipitation with polyethylene glycol and salt, followed by ultracentrifugation in an equilibrium-density cesium chloride gradient. This generates a tube of clear solution that is most dense at the bottom of the tube and least dense at the top; phage particles migrate to form a visible band of phage particles that have migrated to a predictable region (Fig. 6.4). Thus particles starting higher than this density will sink to the corresponding position, whereas those below will float upward. This concentrated layer
6.7 PHIRE: A Ten-Step Program

Fig. 6.4 Equilibrium density gradient purification of mycobacteriophage particles. A phage lysate harvested from 30 large Petri dishes was collected and centrifuged in a cesium chloride (CsCl) equilibrium density gradient for 16 h at 38,000 rpm. The starting CsCl density is 1.5 and during centrifugation a density gradient is established, with the bottom of tube more dense than the top. Phage particles migrate to a position corresponding to their own density and are visible as a bluish band (arrow). The total number of particles in the band is typically $\sim 10^{13}$.

of biologically pure phage particles can be readily recovered with a syringe and needle. Phage samples in cesium chloride usually store well, although DNA extraction, electron microscopy, and other analyses require removal of the salt from the sample by dialysis.

6.7.4 Step 4: Electron Microscopy

Examination of the newly discovered phage by electron microscopy can be performed at any point after its purification, but it is an activity that students often look forward to and which may have significant impact on them. Up to this point, the idea of having a new phage in hand is something of an abstraction; they can deduce that there must be particles in the tube, because there is something there that kills
bacterial cells and forms plaques. But the tube just has the appearance of a clear liquid.

Visualization of the virus particles changes and clarifies the nature of the phage (see Fig. 6.2). Students specifically comment on the impact of this visualization in their notebooks – it is as though a veil has lifted and the nature of the beast is revealed (Hanauer, 2007). A picture is worth a thousands words and provides a concrete description of an entity that could only be assumed to exist through an interpretation of its behavior. This is not only satisfying for the students, but raises their confidence in the interpretive aspects of science employed in the previous steps.

### 6.7.5 Step 5: Nucleic Acid Extraction and Restriction Analysis

Nucleic acid is extracted from the particles by chemical treatments that solubilize or destroy the integrity of the protein coat of the viruses. In our experience with the mycobacteriophages, all of the phages we have isolated contain DNA, although RNA phages (which require somewhat different procedures) are found with other hosts. Many different methods have been described in the literature, although we typically use a mixture of phenol and chloroform, which is immiscible with the phage buffer and thus easily removed. The DNA is soluble in the phage buffer and can be collected by precipitation with ethanol, washed, and resuspended in a suitable storage buffer.

Once the DNA is isolated it is helpful to determine the concentration of DNA and to perform a preliminary analysis by restriction digestion and gel electrophoresis. There are two alternative methods for determining the DNA concentration. First, the optical density at a wavelength of 260 nm can be measured in a spectrophotometer, and using the extinction coefficient of DNA (i.e., its absorption parameters) the concentration can be easily calculated. An alternative method is to examine a DNA sample by electrophoresis (see later) and compare the degree of staining with samples of known concentrations. The first is the most accurate but since other molecules (such as RNA and protein) can also absorb light at this wavelength, it is often useful to perform both methods and compare the results.

Useful and relatively simple insights about a genome can be gained from digesting the DNA with restriction enzymes and analyzing the results by agarose gel electrophoresis. Restriction enzymes are useful because they have the property of cutting the DNA at specific recognition sequences, usually a string of between four and six consecutive DNA bases (see Sect. 6.7.6); because any four-base sequence will occur more frequently than any (for example) six-base sequence, a four-cutter enzyme will cut more often and thus generate more DNA fragments than a six-cutter enzyme. Numerous different restriction enzymes are commercially available and their recognition sites are known.

Restriction digestion is easily performed by incubating a sample of DNA with a restriction enzyme and a simple buffer. Once the reaction is complete, the pattern of fragments generated can be determined by agarose gel electrophoresis, a process by which the DNA fragments can be separated according to their size. A gel matrix
is generated by dissolving agarose (a gelatin-like substance) in a simple buffer by boiling and pouring the solution into a flat mold; a comb containing a series of teeth is then inserted toward one end of the mold and the agarose allowed to cool and set. The comb can then be removed to reveal a series of wells (small square holes) and the gel set in an electrophoresis apparatus to which buffer is added. The opposite ends of the apparatus contain electrodes that can be connected to a power supply that generates an electric current. DNA solutions, each containing different sized fragments, can then be inserted into each of the wells and the electric current applied. Since DNA is negatively charged, the fragments migrate within the electric field toward the positive electrode, but the smaller fragments snake through the agarose matrix faster than the larger ones and thus after about an hour or two of electrophoresis, the fragments are well separated. The approximate sizes of the fragments can be readily calculated by comparing their migration to those in a “standard” or a “marker”, which contains a mixture of fragments of known sizes. An example of such an analysis is shown in Fig. 6.5.

For a preliminary analysis where the DNA sequence is not yet known, performing restriction analysis is most useful for comparing a genome with other genomes. One example is the circumstance in which several plaques are recovered on the initial isolation plate, and it is not clear whether they correspond to completely

![Agarose gel electrophoresis of restriction-digested phage DNA. Approximately 0.5 µg of phage Mikdud DNA was digested with restriction enzymes and the resulting DNA fragments separated by agarose gel electrophoresis. The position of the loading wells and the direction of DNA migration are shown; smaller DNA fragments migrate hastily and are toward the bottom of the gel. Lanes contain marker DNA of known sizes (M), uncut phage DNA (lane 2), or phage DNA digested with BamHI, ClaI, EcoRI, HaeIII, HindIII (lanes 2–6 respectively).](image-url)
different phages or whether they are related. Comparison of the restriction patterns will address this question, since if they are closely related then the restriction patterns will be similar.

A second but equally important utility of this analysis is for comparing the restriction pattern determined for a particular enzyme with that predicted once the complete genome sequence is determined. Because restriction enzymes cut at known specific sequences, the precise pattern can be predicted by computational analysis. Although there are some exceptions, for the most part a close correlation between the predicted and empirically determined digestion patterns is observed. The process of sequence determination involves breaking the genome into smaller pieces and then reassembling them like a jigsaw puzzle into the big picture of the complete genome sequence. Occasionally this assembly can go computationally awry, and this can be revealed through non-congruence of the predicted and empirically measured patterns.

6.7.6 Step 6: DNA Sequencing

To determine the genome sequence of the phages, we have primarily used methods that involve shotgun library construction followed by automated DNA sequencing using an ABI3730 sequencer. This is technically more complex than most of the prior steps. Our laboratory in Pittsburgh has considerable experience with this, so we closely guide students through the process. The standard process requires breaking of the DNA into smaller pieces (2–3 kbp) by hydrodynamic shearing, generating a library of plasmid clones in which each plasmid molecule contains a different and randomly selected segment of the phage genome, performing sequencing reactions in a thermocycler, and analyzing in the automated sequencer. The number of plasmid clones needed varies, but is typically 250–500 and depends on the size of the genome and the quality of the library. For the mycobacteriophages, the average genome length is about 70 kbp, but ranges from 40 to 150 kbp.

Each of the sequencing “runs” usually generates about 800 bp of sequence information. In brief, DNA is a polymer of repeated units, and genetic information is stored in the order, or sequence, of these units. There are typically just four types of units (nucleotides) containing four different bases – adenine, guanine, cytosine, and thymine, abbreviated as A, G, C, and T, respectively. In dsDNA, two strands of nucleic acid are held together through pairing of these bases, with the strict rule that A pairs with T and G pairs with C. The sequence of the bases is different on the two strands, but because of the base-pairing rules, if the sequence of one strand is known, then the sequence of the other strand can be predicted. Thus in a typical shotgun sequencing project each “run” gives a sequence of about 800 As, Gs, Cs, and Ts, and about 500–1,000 individual “runs” are needed (usually by determining the sequences from both ends of each plasmid clone). Since the phage pieces are randomly cloned and sequenced, the overlapping pieces can be computationally assembled together, such that a contiguous sequence of bases corresponding to the entire genome can be constructed.
Once the shotgun part of the project is completed, the sequence can usually be computationally assembled into one or a small number of “contigs” of continuous sequence information. Going from this point to a finally completed sequence can be tricky and requires the confirmation of weak areas in the sequence information, joining of the contigs, ensuring that information is acquired on both DNA strands throughout, and precise determination of the phage genome termini, if it contains defined ends.

In the past several years there have been dramatic advances in DNA sequencing technologies, especially with the pyrosequencing approaches as employed by the 454 sequencer (Roche Inc.) and the methods used by the Solexa genome analyzer (Illumina Inc.). These high-throughput methods can be applied to phage genome sequencing and should greatly simplify sequence acquisition especially if multiple genomes can be determined at the same time. The potential downside is the removal of much of the sequence determination methodology out of the hands of the student-researchers into a core facility. However, the gains in time should be considerable and 454 sequencing is anticipated to ensure that a higher proportion of students achieve complete genome sequencing and can move onto subsequent analyses.

6.7.7 Step 7: Genome Annotation

The raw DNA sequence information holds the informational code with instructions for how the phage grows and survives. Bioinformatic analysis involves the identification of genes and other features through computational analyses, using well-established sequence patterns to predict the overall genome organization. We have made use of the DNA Master program (J. Lawrence, University of Pittsburgh) to facilitate this process, enabling rapid and accurate predictions of gene locations. While this can be accomplished in just a few seconds of computer time, we recommend an up-close manual inspection by students of all of these data. This may take a good deal longer, but it is a powerful way to familiarize students with the nitty-gritty essence of molecular biology, for instance, how the two DNA strands relate to each other, the correspondence of DNA and amino acid sequence information, the difference between transcription and translation, and where genes start and stop. Usually this refinement of the initial crude annotation results in several revisions of gene parameters, identification of additional putative genes, and avoidance of genes improperly predicted by the automated methods. A screen shot of how the information appears to the student is shown in Fig. 6.6.

6.7.8 Step 8: Comparison of the DNA Sequence to Known Phage Genomes

With an annotated genome sequence available, the real fun begins, as the genome can be compared with other known genomes to explore the differences and similarities. There are fundamentally two levels of comparison that can be performed. First,
Fig. 6.6 DNAMaster: A program for phage genome annotation. Screen shots are shown from several representative windows of DNAMaster (A–E). (A) Window presenting features of mycobacteriophage BPs. The “Feature” tab is selected, showing a list of the genes with their coordinates. (B) The “Compare Genomes” window showing the results of a BlastP search of phage BPs gp7. (C) Comparison of phage BPs and phage Halo genomes, with each yellow box indicating the position of genes that are shared between the two genomes. (D) The “Frames” window showing the positions of stop and start codons in all six reading frames. The positions of identified opening reading frames are colored either green (forward direction) or red (reverse direction). (E) Plot of base composition showing the average GC% content across the BPs genome. These represent just a small subset of the available functions with DNAMaster (available as a free download from http://cobamide2.bio.pitt.edu) (see Color Insert)

the entire genome can be compared at the DNA sequence level with other known genomes. It is very rare to find substantial or significant DNA sequence similarity between phages that infect different bacterial hosts, and even phages infecting a common host can be totally different to each other. There are several readily available computer programs (such as Dotter or Gepard) for DNA sequence comparison. If several genomes are to be compared at the same time, the individual sequences can be concatenated and then compared. A typical example is shown in Fig. 6.7, in which dots are placed on the display where two sequences show similarity (specific parameters influencing the stringency of the comparison can be chosen). A diagonal line appears showing that each genome is the same as itself, but any lines off the diagonal reflect similarities between different phage genomes.

Comparison of the nucleotide sequences of 60 mycobacteriophage genomes reveals two fundamental facets of these phage genomes. The first is that even though these phages can all infect the same host (i.e., *M. smegmatis*) and are therefore in direct genetic communication with each other, there is a high degree of genetic diversity; they are all clearly not relatively minor versions of a common sequence. The second is that the relationships are not homogeneous, and there are groups
Fig. 6.7  Nucleotide sequence comparisons of phage genomes. (A) The nucleotide sequences of mycobacteriophages Che9c and Brujita were compared using the program Gepard, using a sliding window of ten nucleotides. Diagonal lines indicate that the genomes share significant nucleotide sequence similarity. (B) Nucleotide sequence comparison of all 60 sequenced mycobacteriophage genomes using Gepard as in (A). The names of the clusters are shown above.

or clusters of genomes that appear to be more similar to each other, than to the other phages. Of the 60 phages, 55 can be clustered with at least one other comprising nine clusters (A–I). It should be noted though that the inclusion of a phage genome within a cluster is largely arbitrary, and the closeness of the relationships is determined by the appearance of the nucleotide comparison in plots such as that shown in Fig. 6.7. It should also be noted that some genomes can have nucleotide sequence similarity to more than the cluster, but spanning just a relatively small genome segment. This clustering is thus one of representational convenience, rather than one reflecting taxonomic differentiation (Hatfull et al., 2006). The appearance of clusters has changed dramatically as additional genomes have been added to the comparison and is expected to continue to do so. We predict that with greater coverage of the mycobacteriophage population the boundaries around these clusters will become fuzzy and may largely disappear in favor of a continuum of relationships.

6.7.9  Step 9: Comparative Genomic Analysis

The average size of bacteriophage genes is smaller than that of bacterial genes and is typically 600–700 bp (Hatfull et al., 2006; Pedulla et al., 2003); thus in a 70 kbp genome there are usually 100–110 genes. A key question is “what do these genes do?” Although this may require detailed experimental studies, much can be learned by comparing a new genome sequence to known phage genomes at the DNA sequence level and the amino acid sequence level. There are commonly
available Internet tools such as BLASTP (for the amino acid sequence comparison), which can be used in conjunction with the tools and databases available at the National Center for Biotechnology Information (NCBI) site; the information gained by comparing predicted genes of a “new” genome to all known genes is extremely valuable.

When comparing a genome of – for example – a mycobacteriophage, there is a greater likelihood that the predicted gene products will have similarity (and thus descended from a common ancestor) to other mycobacteriophage gene products than to anything else in nature. However, we know now that the genome architecture of phages is typically mosaic, with different segments or modules of the genome having different evolutionary histories. This mosaicism can be best represented by examining in detail how each of the predicted genes relates to other mycobacteriophage genes. Steve Cresawn (James Madison University) has devised a program “Phamerator” that performs pairwise comparisons of all mycobacteriophage gene products and assorted then into phamilies (Phams) of related sequences. Phamerator can then display the genomes with each of the genes annotated in accord with its relatives (Fig. 6.9), providing an especially informative comparative genome representation demonstrating the mosaic relationships. The phamerator also produces individual bacteriophage genome maps (Fig. 6.8). Phamerator can also readily provide a representation of the relationships of each gene within a given phamily, which we refer to as “phamily circles” (Fig. 6.9). These phamily circles are essentially phylogenetic representations, but differ fundamentally from the more traditional diagrams of phylogenetic trees. One key advantage is that phamily circles enable

Fig. 6.8  Genome map of mycobacteriophage L5. The L5 genome is represented as a horizontal bar with markers spaced at 1 kbp intervals. Each of the predicted genes is shown as a colored box either above or below the genome; those above are transcribed rightward, and those below are transcribed leftward. The map was generated by the program Phamerator, which sorts the predicted genes into phamilies (Phams) as functions of their predicted amino acid sequence relatedness. The Pham number is shown above each gene with the number of Pham members in parentheses, and each gene is colored according to its Pham. Some of the predicted gene functions are noted (see Color Insert)
Fig. 6.9 Phamily circles. Phamily circles for Pham38 and Pham41 are shown. In the L5 genome (see Fig. 6.8), genes 58 and 59 are members of these two Phams, respectively, and the phamily circles suggest that they have distinct and different evolutionary histories. Each of the phages is listed around the circumference of each circle and an arc is drawn between members of each phamily with line thickness corresponding to strength of the relationship (see Color Insert).

the inclusion of all bacteriophage genomes in the comparison, even those that do not have a member of that particular phamily. Thus the phamily circles can be compared to reveal not only the strength of the relationships, but who participates in them.

This comparative genome analysis can get quite complicated, but it provides a wonderful forum for thinking about genome evolution and for understanding phylogenetic analysis. Performing these comparative analyses provides a really terrific opportunity for students to gain a first-hand experience in what evolution is and how it works. For example, if two adjacent genes in a genome are members of phamilies that have quite different phylogenies (see Fig. 6.9), then at some point in their evolution one or more recombination events must have occurred between them. A simple explanation for how this occurs is via an illegitimate – effectively random in regard to the specific DNA sequence – recombination event at some point in their history. When and how this occurs remains unclear, but the concept of genetic variation occurring in the absence of design or direction, followed by natural selection for the successful progeny, is a very nice illustration of the central tenets of Darwin’s model for the origin of species. Interestingly, because bacteriophage genomes must also conform to fairly strict size constraints – i.e., there must be an appropriate length of DNA, because either too little or too much will not stably fit into the phage head – new variants could at least temporarily be selected other than on the basis of their function, in a non-Darwinian manner. Thus non-functional gene variants could survive within the phage population, but serve as substrates for subsequent recombination events giving rise to yet more new, possibly functional, variants. Genetic mosaicism thus represents a rather rich view of the process of biological evolution.
6.7.10 Step 10: Publication

An indication of the relevance of the genome sequence information is reflected in the importance of its publication. This is typically in one of two forms. First, all of the genomes that are completely sequenced and annotated are submitted for inclusion in one of the commonly available sequence databases (such as GenBank), from which they are accessible to the scientific community. This includes the authors of the work as well as all of the phage genome details.

The second form is the publication of the genome sequence information in the peer-reviewed scientific literature. When and how this occurs for each genome is dependent upon what has been learned from its isolation and characterization. Typically, submitting a manuscript for publication that contains only a single phage genome sequence requires either a stunning finding that warrants publication or considerable additional analysis and experimental dissection (see later). Much of what we have learned from phage genomics in the past few years has emerged from comparative analyses, and students who have contributed substantially to the work have been included as co-authors (Pedulla et al., 2003). In one such publication (Hatfull et al., 2006) a specific description is provided as to what each of the student-authors contributed.

6.8 Is That All?

It is helpful to understand that while learning about the genetic diversity of the bacteriophage population is important, the genomes and the genomic information also represent a resource and a good starting point for many other scientific inquiries. The reason why we have primarily focused on the viruses that infect mycobacterial hosts is not because we anticipate them to be systematically different to phages of other bacteria, but because the hosts are important and interesting in their own rights. *Mycobacterium tuberculosis* is a particularly special pathogen and kills nearly two million people each year across the globe (Raviglione, 2003). Approximately one-third of the entire world’s population is infected with *M. tuberculosis* and the emergence of especially nasty drug-resistant strains further compounds the problem. In a 2006 study of an outbreak of extensively drug-resistant tuberculosis (XDR-TB) in Tugela Ferry, South Africa, 52 of the 53 patients died and the median time to death from collection of an initial sputum sample was 16 days (Gandhi et al., 2006). There is a desperate need for effective vaccines, new drugs, and rapid diagnostic tests for determining drug susceptibility. From the molecular microbiologist’s perspective, these advances are likely to depend on having a strong understanding of the biology of *M. tuberculosis* and the availability of the genetic tools required for its facile genetic manipulation.

Because of the slow growth of *M. tuberculosis* (doubling time of 24 h) and its pathogenicity, the development of a system for genetic manipulation has been
6.9 What’s in a Name?

Scientific discovery is a powerful force. Within the context of the PHIRE program, discovering a new virus in the environment is an intrinsically exciting activity and is intimately associated with the concept of project ownership. Students recognize that isolating a new virus may be of importance to the community, but it is also important for them personally. If they do not decide to pursue its characterization and unlock its secrets, then who will? The discovery they have made is “money-in-the-bank,” and further studies will result from the proud investment of their time and energy.

This sense of project ownership and discovery is further reinforced by the key step of naming the phage that they have discovered. Students are encouraged to use individual names rather than any more systematic scheme, with the justification that our genomic understanding of bacteriophages is that they cannot be readily arranged into hierarchical lineages because of the extensive amount of horizontal genetic exchange that dominates their evolution. The term “species” cannot be easily applied to bacteriophages, no matter what definition of the term is applied (Lawrence et al., 2002). So we use names that reflect this individualistic character. Examples of some of the mycobacteriophage names include Corndog, Barnyard, Giles, Pipefish, and Rosebush.

It is not uncommon for students to take some time in making the decision as to what to name their phages. A particularly illustrative example occurred when one of us (GFH) was giving a presentation of the PHIRE program and happened to have a program student in the audience. The student was asked to come to the podium (without prior prompting) and talk about his involvement in the PHIRE program. In response to the question “Why did it take so long to decide on the choice of a name for your phage,” he thought for a while and then responded “Because it is so important!” This is a fairly common reaction, the idea that the phage they have isolated and characterized becomes a part of the scientific landscape and one that may be used by others for many years to come in various ways. Its name is therefore important, and not something to be taken too lightly. Enabling students...
with the authority to provide a name (we usually hold the power of veto, just in case) couples closely with the concept of project ownership, further cementing the association of the student with the phage.

6.10 Who’s Qualified for the PHIRE Program?

The PHIRE program unusually combines the idea of being able to contribute to key scientific advances without the need for extensive and specific training or experience. In essence, this is a program that anyone – literally anyone – can participate in. As noted above, the techniques required in the initial stages are quite simple and can readily be performed in most high-school settings. There is also not really much that you need to know that would prevent someone from doing these initial steps. This should not be confused with the notion that there is no scientific content to the PHIRE program; clearly the opposite is true, but engagement and success in performing the experiments provides a platform to learn about the ecology, microbiology, genetics, evolution, and biochemistry that emerge within the program. Participants of the PHIRE program include undergraduates, high-school students, teachers, and even middle-school students. While not everyone will necessarily be fully engaged for all parts of the program or be inspired to complete all the steps described above, we have yet to find someone who is unqualified to start in the PHIRE program.

Perhaps this low barrier to participation seems unimportant or trivial, but it is a truly key element of PHIRE. Most importantly, all students have an equal opportunity for participation, and there is no need to employ biased and inappropriate conditions for selection. For example, when choosing students for participation we do not give undue weight to standardized test scores because of the limited validity of these tests for our educational program that is organized around the performance of actual scientific research. An expression of curiosity by students is usually a good indicator that they will benefit from the program, and we generally aim for as diverse a group of students as is reasonably possible. In essence, everyone is qualified for the PHIRE program, and it provides opportunities for students to find out if this is an endeavor in which they excel even if their past experiences may not necessarily have predicted it.

6.11 Parallel Projects: Pros and Cons

The fact that we are able to describe a ten-step project plan as part of a reflective process of program development and active assessment underscores another core PHIRE component. Essentially, while students enjoy the novelty of project ownership and discovery, the processes and techniques employed are essentially the same for each student. This can therefore be thought of as a massively parallel project
structure and one that comes not only with some notable attributes but also with some detriments.

The detriments are reasonably obvious. If students are engaged in parallel projects then there is a compromise of the concept of doing truly independent open-ended research, where the outcome is not known and there are many decision points along the way. We regard this as an unfortunate compromise, but it also needs to be recognized that inclusion of any more than a very small number of students is difficult if each has a truly independent project. To conduct more open-ended projects students typically need a more sophisticated knowledge of the subject matter and usually require specific and close supervision. Unfortunately, the ratio of supervisors (graduate students and postdoctoral fellows) to student-researchers rarely supports more than a small handful of students performing open-ended projects in any medium-sized research laboratory.

There are, however, considerable attributes to having parallel projects. Most importantly, there is a well-defined set of approaches that will be used and these are established and thoroughly streamlined. But the supervision problem largely disappears, since more experienced PHIRE students are well positioned to assist more junior PHIRE students, having already mastered most of the early phases. This peer supervision also releases the constraints on the number of students in the program. In fact having a greater number of students is an administrative asset. It decentralizes the program operation and enhances communication between PHIRE participants. Thus, while acknowledging the compromises, there are really important practical advantages. Furthermore, students are nicely trained for moving onto more independent projects as they emerge from their PHIRE successes.

6.12 Mentees and Mentors: Opportunities and Responsibilities

Students entering the PHIRE program are usually assigned a specific mentor, another student who is already within the program. In most cases this will correspond to a more “senior” undergraduate mentoring a more “junior” undergraduate or an undergraduate mentoring a high-school student, although the apparent academic hierarchy is merely superficial. In several instances, for example, high-school students have mentored undergraduate students, provided that they are established in the program and been successful through the early phases.

This mentor–mentee system is important for program administration of the parallel projects as described above, but it has a more important underlying role. Students entering the program are the beneficiaries of the opportunities that they are provided with – opportunities to engage in authentic scientific discovery. As mentors, they learn that with these opportunities come the responsibilities associated with passing on what they have learned to others; mentor training can be readily facilitated through the Entering Mentoring program established by Jo Handelsman and colleagues (Handelsman et al., 2005). Mentoring is as central to the PHIRE program as it is possible. It is the “I” in PHIRE, the integrating of research and
education, with the opportunities to do research, inseparably associated with the responsibilities of education. The PHIRE program thus strives not just to introduce students to the world of science – who scientists are, how science is done, and the scientific environment – but places them from the beginning as scientist/educators or educator/scientists. Where divisions between our research and educational missions throughout academia have caused evident damage, this emphasis is an especially positive attribute to the PHIRE program.

6.13 Scheduling and Flexibility

The PHIRE program is not a course, and the time that students invest in their project is mainly determined by them, just as graduate students or postdoctoral researchers do. Students may spend time in the program during the summer months, taking the advantage of large blocks of available time for research, or they may be engaged throughout the academic year. For undergraduate students, finding the time can be challenging, especially early in their programs since the first two years of college/university education can often be demanding with required and core course work. High-school students also have many demands on their time, but several have worked throughout the academic year on a PHIRE project, whenever time was available to them. Clearly, the project moves faster when more time is invested in it, but significant parts of the project can be performed within a flexible schedule, and there are few phases that require large extended periods of intense work that have to be done within a particular schedule at the expense of all other demands.

Once again, it is tempting to consider the suitability of PHIRE projects with flexible scheduling as a relatively minor program component, but it fits well with the other program components. In particular, in the absence of a pre-arranged rigid schedule, students must be self-motivated to engage in the research and schedule their time accordingly. With many students, this motivation is closely linked to the rewards of discovery and project ownership and a quick realization that the successes are often the direct result of the degree of commitment to the project. Without a rigid course-like structure, project ownership, discovery, and research successes are absolute requirements of the program.

6.14 Multiple Milestones of Success

While the PHIRE program involves quite a well-formulated path, we recognize that not all students will succeed with all stages. Students may not have time, some will progress slower or faster than others, various difficulties can be encountered along the way, and some students learn that this endeavor is something that they don’t really have their heart in. Nonetheless, students achieving only part of the process can still enjoy some degree of success. They will with near-certainty have isolated a
new phage, have been able to purify it, and prepare stocks. During this process they have learned about the laboratory environment and the scientific process, worked alongside other professional scientists, and absorbed a substantial amount of microbiology! In an environment within the United States in particular where there are significant concerns about the level of scientific literacy among the general public, these would seem not insignificant accomplishments.

6.15 Not All Failure Is Bad!

Perhaps it is obtuse to laud the attributes of failure, but there are surely some benefits. Being a successful practitioner of scientific research requires many skills, but one of the most important is problem solving – knowing how to recognize and address substantial experimental and technical challenges and how to circumvent or resolve them. Becoming familiar with these problem-solving issues and to address them is an important part of the experience in learning how to do science and provides a sense of empowerment to students when they learn how to solve problems for themselves. Furthermore, it is not uncommon to identify students who are interested and curious in doing research that have a history of extraordinary academic achievement, in which they have always excelled at the classroom challenges. It can be a very humbling experience for such students to find that it is they – rather than the academically “average” student next to them – that has yet to identify a phage.

While some of the steps in phage hunting and genomics are more technically challenging than others, the phage isolation procedure is a playing surface-leveling experience. Success is not something that can be easily controlled because the presence of phage particles for the host you are studying is not readily predictable (other than recognizing that phages tend to be found in the environment in the same places as their hosts). A lack of care in following the procedures, sloppy work, and poor attention to detail can all make the isolation more difficult, but academic preparation, knowledge of microbiology, and supreme self-confidence, all have little influence. Understanding what is within and what is beyond the control of the scientist is an important lesson for all students to learn.

6.16 Transitions: From Concrete Beginnings to Abstract Representation

When practicing science, it is easy to forget that most of what we do is to develop and debate abstract concepts. Learning about science is in large part the art of acquiring the language that is required and being able to understand and manipulate concepts that are often divorced from grounding in our everyday lives. Thus the relationship between abstract concepts and their concrete roots is a critical component in the learning of science.
An appealing aspect of the PHIRE program is that there is a built-in transition from the very concrete beginnings – digging up dirt and collecting it in a tube – to the thoroughly representational process of decoding thousands of letters and numbers of a genome sequence rolling over a computer screen (Hanauer et al., 2006). The transition is fairly gradual, as dirt collection leads to growth of bacterial cultures, followed by preparation and electron microscopy; this leads to DNA isolation and the much more abstract processes of DNA sequence determination and annotation.

6.17 Seven Attributes of the PHIRE Program

In this chapter we have hopefully established an understanding of the PHIRE program, its elements, and the reasons why we believe it has some successes associated with it. It is also helpful to note that it represents a very different perspective to how science is typically taught, especially in the K-12 setting. The K-12 culture – especially at the high-school level – is commonly a process of the communication of our known facts and concepts from the instructor to the students, and students who do well are those that perform well at listening, processing, retaining the information, and subsequently recalling it in the context of standardized and other tests. However, as discussed in Chap. 2, the process of scientific research operates on a different set of principles. It requires challenging our current understanding and utilizing inquiry-based operations to elucidate new knowledge. Doing so requires prioritizing ideas and concepts in order of importance and of thinking about those issues and topics that we don’t comprehend. Finally, it requires an idea as to how new knowledge is obtained and indeed the basis upon which we assert that we know what it is that we know.

The concept behind the PHIRE program is that inquiry-based activities are the only way that students can learn science, as opposed to the facsimile that currently is adopted in most K-12 settings and sometimes at the undergraduate level, especially in the freshman and sophomore years. We recognize that, for all students, to take advantage of such experiences poses major challenges, but we also hope that these experiences with PHIRE and other inquiry-based programs will provide a focus on what needs to be done to make this happen. In this chapter we have discussed the scientific and education attributes of the PHIRE program, but in summary will list seven attributes here that are translatable items and could be used as building blocks to construct other such programs with strong integration of research and education (Hatfull et al., 2006).

6.17.1 Attribute 1: Technical Simplicity with Transition to Complexity

Phage isolation is technically simple, and while some techniques can be more demanding they occur later in the process. Even finding genes begins with the simple concept of pattern recognition.
6.17 Seven Attributes of the PHIRE Program

6.17.2 Attribute 2: Conceptual Simplicity with Transition to Complexity

Some of the concepts that will be learned with the PHIRE program are quite complex, but they are acquired in an inquiry-based context and high academic achievement is not required for starting the program.

6.17.3 Attribute 3: Flexible Scheduling Compatibility

Students have many demands on their time besides doing scientific research, and flexible scheduling capability is important. This approach to scheduling provides a degree of control to the students over the progress of their project.

6.17.4 Attribute 4: Multiple Achievement Milestones

The only serious and significant “failure” within the PHIRE program is the inability to isolate a phage; however, this is rare (about 1% in our experience). All others enjoy some degree of achievement, even if not all steps are accomplished.

6.17.5 Attribute 5: Parallel Projects and Mentoring

There are numerous advantages to having a parallel project structure, of which the most important is an inherent suitability for students to mentor other students.

6.17.6 Attribute 6: Authentic Publishable Research

The science within the PHIRE program is important and interesting to a broader audience. There are therefore rich opportunities for publications in the peer-reviewed literature. Also, students readily recognize the difference between authentic research and an exercise, and the latter is often not well endowed with motivation.

6.17.7 Attribute 7: Project Ownership

The association of the discovery with the student is powerful, and naming the phage that they have discovered provides incentives for continued involvement and commitment to the project.
Chapter 7
The PHIRE Program Assessment Strategy

7.1 Introduction

This chapter describes the specific assessment strategy developed for the educational assessment of the PHIRE program. The PHIRE assessment strategy covers formative diagnostic and summative aims of an education in bacteriophage scientific inquiry. The strategy proposed in this chapter consists of the integration of five different types of tool that cover the four knowledge sources of physical, cognitive, representational, and presentational knowledge and is organized across the scientific inquiry timeline of bacteriophage isolation and genomic identification. The overall approach follows the guidelines of authentic scientific inquiry assessment and designs assessment procedures following an informed understanding of the specific scientific and educational process that is being assessed.

As specified within Chap. 5, the process of designing an assessment strategy and constructing assessment tools consists of a series of stages. As an initial starting point the scientific inquiry process at the heart of the educational program needs to be analyzed and described (see Chap. 6 for a description of the stages of this program); then a series of decisions concerning the aims of the program and which specific representations and their adjacent knowledge are considered to be significant need to be made; following this a strategy and a set of assessment tools can be designed. The design of this chapter follows a similar plan starting with an analysis of aims, a detailed analysis of the decision-making process at significant milestones of the phage-hunting process, and finally the proposition of an overall strategy for assessment of this scientific inquiry process. Further clarification and exemplification of the specific assessment tools used in the PHIRE program appears in the next chapter.

7.2 Aims of Undergraduate Phage Hunting

The aims of undergraduate science education within the PHIRE program have been articulated by the prime researcher and educator of the program, Prof. G. Hatfull. In his Executive Summary of the Phage Hunting Outreach Program (2006), he
states that students should “learn how to do scientific research and how to discover nature’s secrets.” The educational program is designed to “elucidate the mechanism for viral evolution and to determine the extent of viral diversity through the isolation and characterization of novel bacteriophages.” The program not only provides the framework within which each student can “isolate a new bacteriophage” but also has the wider aim of “demystifying the culture of scientific discovery” and provide each student with the scientific “understanding that will enhance whatever career path they ultimately pursue.” Finally the program wishes to “correct commonly held misconceptions about scientific discovery – what it is and who does it – and to show these students that they can become the next generation of scientific explorers.” In terms of assessment, the aims of this program for undergraduate phage hunters can be stated as follows:

1. **Summative Assessment: Development of Substantive Knowledge Concerning Bacteriophage**: As with any other undergraduate science program, the PHIRE program has the aim of developing substantive scientific knowledge in relation to the core scientific concepts that inform and direct the scientific inquiry process. In the case of the PHIRE program these core concepts relate to three basic issues: the nature and mechanisms of viral evolution, the physical and genetic characteristics of bacteriophages, and research into the potential applications of bacteriophages in associated fields such as medicine. Data collected from this form of assessment could be used to evaluate the substantive factual knowledge acquired by students involved in this program.

2. **Summative Assessment: Development of Procedural and Representational Knowledge Concerning the Microbiological Process of Isolating and Genomically Annotating and Analyzing a Novel Bacteriophage**: The undergraduate PHIRE program involves a specific scientific inquiry program and as such is designed to enhance student understandings of the scientific procedures and methods involved in the isolation and characterization of novel bacteriophages and the microbiological aspects of these organisms and their interactions with the environment. This knowledge base includes basic microbiological and bioinformatic epistemology and a clear understanding of the representational aspects of the phage-hunting process. The aim of the assessment procedure is to elicit assessment data that can be used to determine the levels of procedural and representational knowledge development achieved by the participants in this program. These data can be used to describe the outcomes of the educational program in relation to the specific aspects of the scientific inquiry process of the isolation and characterization of bacteriophages.

3. **Summative Assessment: Presentational Knowledge and Self-Perception Concerning the Bacteriophage Scientific Inquiry Process**: The PHIRE program aims to help students to develop a self-concept as a researcher and the professional knowledge associated with this role. This knowledge includes the ability to present research finding within formats that are authentic to the ways in which professional researchers presented their research. This form of knowledge is presentational and consists of genre of presenting information at professional
settings. The data collected from this assessment could be used to evaluate the outcomes of the student’s scientific inquiry process and the ability of the student to clearly conceptualize and present these results in a professional manner.

4. Formative and Diagnostic Assessment – Procedural, Substantive, and Representational Knowledge of the Microbiological Process of Isolating and Genomically Annotating Novel Bacteriophages: The PHIRE program is an educational program designed to enhance scientific learning through undergraduate participation in an authentic scientific inquiry process. The PHIRE program aims to provide every student with the ability to master the procedural, substantive, and representational aspects of the phage-hunting process. Accordingly, formative and diagnostic assessment data are collected to ensure that quality feedback can be provided to the students so that the scientific inquiry process can be completed successfully. As opposed to other forms of educational practice a basic assumption of the PHIRE program is that there is real scientific value in the specific scientific inquiry processes conducted by each student. This educational program functions through a form of directed mentorship in which students conduct guided scientific inquiries, and in this context ways of collecting data that can be used to assess the knowledge of the student and provide direct and informed feedback are useful for enhancing the students’ ability to complete a scientifically significant inquiry process. The basic aim of this assessment procedure is to provide scientist-educators with information to provide quality feedback to students involved in the inquiry process.

7.3 Detailed Analyses of Decision-Making Reasoning and Calculation in the Bacteriophage Isolation and Identification Process

As defined within Chap. 5, decisions need to be made as to what representational stages constitute milestones within the developing scientific inquiry process and can be used for assessment purposes. Once these representational stages have been chosen a full analysis of the adjacent knowledge sources needs to be conducted. In relation to the bacteriophage isolation and identification process there are several significant decisions that have to be made in order to complete the process. The following is the outline of significant decision making and calculation that are made according to the stage in the process:

a. Step 1 – Phage Isolation: Identification of plaques
b. Step 2 – Phage Purification: Generating a pure phage population
c. Step 3 – Phage Amplification: Calculation of lysate concentration: Selection of maximum bacteriophage yield
d. **Step 4 – DNA Comparison**: DNA restriction enzyme digestion and electrophoresis: Comparison of electrophoresis gel and decision as to value of continuation of process

e. **Step 5 – Genome Annotation and Analysis**: Gene Identification – Selection and identification of genes

Each of these decisions marks a point of transition in the whole process and as such is worth considering within the context of significant representational milestones that can be used for assessment of tool development. Below a detailed analysis of each of these decision-making processes is presented. The analysis presents the basic knowledge that is required to make the decision and describes the decision-making process using the formats of a schematic description and a logical inferential structure. The decision-making process for the first four decisions is within the environment of the wet laboratory and the last decision-making process is within the digital environment of a computer program.

### 7.3.1 Phage Isolation – Identification Through Infection

This conceptual reasoning process allows bacteriophages to be identified. The process is based on a basic understanding of microbiology – bacteriophages are viruses whose hosts are bacterial cells. Bacteriophages reproduce by exploiting the metabolism of the host bacteria. The bacteriophage delivers its genome into the cytoplasm of the bacterial host where it uses the cellular mechanisms of the host to reproduce. This typically results in the destruction of the original cell through the extensive reproduction of the bacteriophages and subsequent destruction of adjacent cells as the progeny of that first bacteriophage infects and destroys those bacteria. Accordingly, based on this characteristic of bacteriophages, a process of initial identification can be defined. If a sample potentially containing bacteriophages is placed in the controlled vicinity of a bacterial cell it will either infect or not infect the bacteria. If it does infect the bacteria, there is evidence of the presence of bacteriophages, which is revealed by the presence of a plaque, a circular zone of cell death causing a cleared zone in a bacterial lawn (see Fig. 6.2). If it does not infect the bacteria (i.e., no plaques are present), we assume that no bacteriophages specific to that bacterial host are present (and capable of infection under these controlled laboratory conditions). Schematically this process can be defined as follows:

**Required Determination**: Presence (+) of bacteriophages (or PFU, plaque-forming unit) in sample

**Sample**: 2 options:
- + presence of plaques (bacteriophages)
- – absence of plaques (bacteriophages)

**Empirical Test**: Samples plated with bacteria

**Results and Analysis**: Infection (presence of plaques) = + bacteriophages
- No infection (no plaques) = – bacteriophages
As a logical structure this can be represented as follows:

a. Bacteriophages infect bacteria
b. If plaques are present in a controlled environment, then bacteriophages are present
c. If plaques are not present, then bacteriophages are not present (or cannot be detected)

7.3.2 Phage Purification Through Selective Infection

This conceptual process allows a single bacteriophage to be selected from a sample which potentially includes several different bacteriophages. A basic assumption of this conceptual process is that during the process of infection, bacteriophages will create plaques that can be visually defined as “characteristic” of that bacteriophage. Accordingly, two different bacteriophages in the same sample may produce two visually different types of plaque morphotypes. If this occurs, then the aim is to generate a sample that contains only a single bacteriophage type. Thus at each stage of infection, when multiple different plaques are found, a choice is made as to which plaque will be chosen for further infection. This selection process gradually purifies the sample to the point at which all the plaques on the plate have the same visual characteristics. This situation is understood as the presence of a single bacteriophage type in that sample. Schematically this process can be represented as follows:

**Required Determination:** Presence of a single bacteriophage type in a sample

**Sample:** 3 options:
- absence of bacteriophages
- presence of a single bacteriophage type
- presence of multiple bacteriophage types

**Empirical Test:** Repeated plating of samples with bacteria

**Results and Analysis:** Visual analysis of morphology of plaques

- Multiple plaques of different size, shape, and appearance = multiple bacteriophages (choose one plaque to repeat test)
- Multiple plaques of same size, shape, and appearance = single bacteriophage type (objective achieved, progress to next stage)
- No plaques = no bacteriophages

As a logical structure this can be represented as follows:

a. When infecting bacteria in a plated sample, each bacteriophage type creates a characteristic plaque according to its size, shape, and appearance.
b. If a plate has multiple plaques of different sizes, shapes, and appearances then different bacteriophages are present in the plate.
c. If a plate has multiple plaques of the same size, shape, and appearance then it is more likely that a single bacteriophage type is present.
7.3.3 **Phage Amplification – Measuring Bacteriophage Concentration to Determine and Produce Maximum Bacteriophage Yield**

The conceptual process involves working out the concentration of bacteriophage within a sample so that a decision on maximum phage yield levels for the production of large samples of bacteriophage can be made. This process involves the evaluation of a concentration (or amount) rather than a dichotomous present/absent judgment. The overall judgment as to the maximum phage yield level of bacteriophage concentration for amplification is done in three stages: first, the lysate is diluted at 10-fold intervals, plated, and titered (measured); second, the titer is manipulated to theoretically produce the maximum phage yield based on the assumption that 10,000 ± plaques will produce that yield, manifested as the desired “web pattern” in which just enough phage particles have been added to infect all – or nearly all – cells that are present; third, the assumed maximum phage yield calculations that will produce the web are tested at five close concentration levels to allow for variations among phage types. An empirical visual analysis of the web pattern on the plate determines the best concentration of phage to be used in the 30-plate infection.

**Required Determination:** Concentration (titer) of bacteriophage in a sample  
**Sample:** 10 levels of dilution for bacteriophage sample  
**Empirical Test:** 10 dilution levels plated  
**Results and Analysis:** Visual analysis of the number of plaques at the most countable dilution and mathematical calculation to determine the titer (PFU/ml)

As a logical structure this can be represented as follows:

a. Each 10-fold dilution of sample produces a 10-fold decrease in the number of PFUs counted.

b. A plate of 20–200 PFUs is counted (this plate is used because it is considered the most countable plate).

c. The number of PFUs counted is used to calculate the titer according to the following equation: PFUs counted/amount of sample plated (µl) × Dilution Factor × 10³ µl/1ml = Titer PFU/ml.

**Required Determination:** Maximum phage yield concentration to be used for 30-plate infection, typically calculated to 10,000 PFU/ml  
**Sample:** Using the titer, the maximum phage yield dilution is calculated, four additional dilutions are plated to empirically identify the desired web pattern:  
- Calculated maximum phage yield (MPY) concentration  
- One magnitude above MPY  
- Half of one magnitude above MPY  
- One magnitude below MPY  
- Half of one magnitude below MPY  
**Empirical Test:** Five dilution levels of lysate plated
Results and Analysis: Visual analysis of plate plaque patterns

- Exhibits required (+) web pattern – determine this dilution level as optimal
- Exhibits a different pattern from the required (−) web pattern – determine that this dilution level is not optimal

As a logical structure this can be represented as follows:

a. A plate with maximum phage yield produces a distinctive pattern that can be recognized.

b. If the plate (each with different dilution levels) has the distinctive pattern of a MPY then that represents the best choice of dilution level for the production of large quantities of bacteriophage particles.

c. If the plate does not have the distinctive pattern then it is not a good choice for the production of large quantities of bacteriophage particles.

7.3.4 Electrophoresis

The conceptual process consists of determining the uniqueness of a bacteriophage DNA sample. This determination is made in relation to the procedures of DNA restriction enzyme digestion (RED) and electrophoresis (see Fig. 6.5). Restriction enzyme digestion of DNA is a process in which a restriction enzyme recognizes a specific sequence of bases and cuts the DNA at those specific sites. The segments generated are potentially of different lengths and each genome will produce a characteristic pattern of DNA fragments. These can be separated and analyzed by the process of electrophoresis in which samples containing mixtures of DNA fragments are loaded onto an agarose gel, to which an electric current is applied. DNA fragments migrate toward the negative pole with smaller fragments migrating faster through the gel matrix than larger ones. Within a given time frame all of the DNA fragments can be separated and their sizes measured by comparison to known size markers. In the determination of the uniqueness of a bacteriophage DNA sample, a series of five (or more) DNA enzyme digests are used individually, and the pattern generated compared with the patterns of previously characterized bacteriophages.

Required Determination: The uniqueness of the isolated bacteriophage
Sample: Two options: Unique pattern of DNA restriction fragments
Identical pattern of DNA restriction fragments

Empirical Test: Electrophoresis gels of restriction enzyme-digested DNA

Results and Analysis: Visual analysis of electrophoresis gels (the visual analysis involves an evaluation of the number, location, and intensity of the bands for each enzyme and the overall pattern of all of the digests, sometimes referred to as its “DNA fingerprint”).

(−) Exhibits unique pattern of DNA restriction fragments
(+ ) Exhibits similarities to a DNA restriction pattern of another phage.
As a logical structure:

a. The electrophoresis of a DNA restriction enzyme digest produces a pattern that is made up of digested DNA segments ordered according to size (as seen according to the number, position, and intensity of the bands).
b. Each bacteriophage DNA will produce a characteristic pattern of digested DNA segments on an electrophoresis gel.
c. If the DNA restriction enzyme digest produces a pattern that is identical to another DNA restriction enzyme digest then the assumption is that this is a duplicate (or VERY similar) of a bacteriophage that has already been identified.
d. If the DNA restriction enzyme digest produces a pattern that is different from the other DNA restriction enzyme digest then the assumption is that this is a previously undiscovered bacteriophage.

*Required Determination:* Evaluation of restriction enzyme digest pattern of a bacteriophage

*Sample:* Electrophoresis of DNA restriction fragments of a bacteriophage DNA sample

*Empirical Test:* Set up digests and run gel electrophoresis

*Results and Analysis:* Comparison of the overall pattern of bands with the digests of other known phages determines the uniqueness of each bacteriophage.

### 7.3.5 Gene Determination (Annotation)

The gene determination process aims to identify genetic sequences that have high coding potential to be genes. This process is interpretive and utilizes a series of sources of information. The basic determination is in relation to the computer file output designating the complete genetic sequence of the bacteriophage DNA sample. The sources of information used in the determination are the DNA Master Program, the GeneMark program, the Glimmer program, the NCBI Gene Blast program, and general principles of bacteriophage genomic structure. The actual process of identifying and recording the identified genes is done in the DNA Master program. Following is a schematic description of each program and its output:

- **GeneMark Program:** Provides a prediction of gene locations according to DNA sequence characteristics that are in common to those of known genes. Visual analysis of a computer output containing a trace depicting the degree of coding potential in each of all six possible frames across the entire genome (Fig. 7.1). In addition to coding potential the output also includes the positions of potential gene start sites (as up ticks) and stop codons (down ticks). An open reading frame (ORF) – i.e., a region of DNA in which stop codons are absent – that contains a potential start site at the beginning, a stop codon at its end, and high coding potential throughout has high probability for corresponding to a gene.

- **Glimmer:** Provides a prediction of gene locations by analyzing the DNA sequence searching for characteristics that correspond closely to internal open
Fig. 7.1  Graphic Output of GeneMark Program. A segment of phage Adjutor genome sequence was analyzed using the program GeneMark, which generates a graph of the coding potential in each of the six possible open reading frames (ORFs) (labeled 1–6; frames 1–3 are expressed in the rightward direction, and 4–6 in the reverse direction). ORFs are identified as *horizontal lines* between translation stop codons (*down ticks*), and potential start codons are shown as *up ticks*. Coding potential is indicated by the line graph. The thick horizontal gray lines (still italicized) are considered areas of interest. One such gene is highlighted in *green* (see Color Insert).

reading frame features. The output is a computer file that depicts the coordinates of the start and stop of each gene (Fig. 7.2).

- **NCBI Gene BLASTing**: Compares the amino acid translation of a potential gene to a database of known protein sequences. The output is a computer file that identifies genes products that are statistically similar to those in other organisms (Fig. 7.3). The file provides a probability score for the quality of the relationship between the different gene products. If the BLASTed gene product has a BLAST score of acceptable significance, this provides further evidence for the gene identification.

- **DNA Master**: Compiles all data outputs together for evaluation by importing GeneMark and Glimmer data, providing a Shine–Dalgarno score to evaluate the potential for different start codons at the beginning of genes, and providing a six-frame translation for evaluation. It is in this DNAMaster program that all parameters and guidelines for the gene identification decision are synthesized into the best evaluation of genes in a genome (Fig. 7.4).
**General Principles of Bacteriophage Genomics:** Phage genomics is an emerging field. The data that guide the decision-making process of “correct” gene identification are fraught with incomplete information. Allowing Glimmer and GeneMark data to find the genes is relatively straightforward and reliable, but identifying the best start codon position involves subtle nuances requiring a trained human interpretive process. When choosing the “best” start codon, the following list of factors needs to be considered:

**Fig. 7.2** Glimmer Program Output. A segment of the phage Adjutor genome sequence was analyzed by Glimmer generating a list of potential ORFs, their start and stop codons, reading frame, and evaluative score. The red box designates one specific candidate gene (see Color Insert)

<table>
<thead>
<tr>
<th>orfID</th>
<th>start</th>
<th>end</th>
<th>frame</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;Adjutor</td>
<td>43</td>
<td>309</td>
<td>+1</td>
<td>5.50</td>
</tr>
<tr>
<td>orf00002</td>
<td>421</td>
<td>1524</td>
<td>+1</td>
<td>11.60</td>
</tr>
<tr>
<td>orf00003</td>
<td>1568</td>
<td>1762</td>
<td>+2</td>
<td>7.44</td>
</tr>
<tr>
<td>orf00004</td>
<td>1772</td>
<td>1936</td>
<td>+2</td>
<td>19.55</td>
</tr>
<tr>
<td>orf00005</td>
<td>1946</td>
<td>2125</td>
<td>+2</td>
<td>2.27</td>
</tr>
<tr>
<td>orf00006</td>
<td>2136</td>
<td>2776</td>
<td>+3</td>
<td>9.08</td>
</tr>
<tr>
<td>orf00008</td>
<td>3773</td>
<td>4543</td>
<td>+2</td>
<td>7.48</td>
</tr>
<tr>
<td>orf00009</td>
<td>4540</td>
<td>6306</td>
<td>+1</td>
<td>13.68</td>
</tr>
<tr>
<td>orf00011</td>
<td>6328</td>
<td>7914</td>
<td>+1</td>
<td>15.10</td>
</tr>
<tr>
<td>orf00012</td>
<td>7925</td>
<td>8137</td>
<td>+2</td>
<td>8.51</td>
</tr>
<tr>
<td>orf00013</td>
<td>8134</td>
<td>8586</td>
<td>+1</td>
<td>6.27</td>
</tr>
<tr>
<td>orf00014</td>
<td>8555</td>
<td>8734</td>
<td>+2</td>
<td>9.56</td>
</tr>
<tr>
<td>orf00015</td>
<td>8822</td>
<td>9067</td>
<td>+2</td>
<td>1.62</td>
</tr>
<tr>
<td>orf00016</td>
<td>9235</td>
<td>9636</td>
<td>+1</td>
<td>6.98</td>
</tr>
<tr>
<td>orf00017</td>
<td>9669</td>
<td>10376</td>
<td>+3</td>
<td>10.88</td>
</tr>
<tr>
<td>orf00019</td>
<td>10392</td>
<td>10811</td>
<td>+3</td>
<td>10.24</td>
</tr>
<tr>
<td>orf00020</td>
<td>10848</td>
<td>12032</td>
<td>+3</td>
<td>12.62</td>
</tr>
<tr>
<td>orf00022</td>
<td>12029</td>
<td>12397</td>
<td>+2</td>
<td>12.81</td>
</tr>
<tr>
<td>orf00023</td>
<td>12400</td>
<td>12732</td>
<td>+1</td>
<td>7.05</td>
</tr>
<tr>
<td>orf00024</td>
<td>12729</td>
<td>13136</td>
<td>+3</td>
<td>7.49</td>
</tr>
<tr>
<td>orf00025</td>
<td>13129</td>
<td>13491</td>
<td>+1</td>
<td>8.23</td>
</tr>
<tr>
<td>orf00026</td>
<td>13488</td>
<td>13907</td>
<td>+3</td>
<td>5.12</td>
</tr>
<tr>
<td>orf00027</td>
<td>13944</td>
<td>14300</td>
<td>+3</td>
<td>8.89</td>
</tr>
<tr>
<td>orf00028</td>
<td>14312</td>
<td>14512</td>
<td>+2</td>
<td>4.14</td>
</tr>
<tr>
<td>orf00029</td>
<td>14541</td>
<td>15344</td>
<td>+3</td>
<td>9.78</td>
</tr>
<tr>
<td>orf00030</td>
<td>15426</td>
<td>16079</td>
<td>+3</td>
<td>10.30</td>
</tr>
<tr>
<td>orf00032</td>
<td>16102</td>
<td>16320</td>
<td>+1</td>
<td>8.00</td>
</tr>
<tr>
<td>orf00034</td>
<td>16322</td>
<td>21094</td>
<td>+2</td>
<td>10.80</td>
</tr>
<tr>
<td>orf00036</td>
<td>21084</td>
<td>22604</td>
<td>+3</td>
<td>14.22</td>
</tr>
<tr>
<td>orf00038</td>
<td>22601</td>
<td>24256</td>
<td>+2</td>
<td>12.44</td>
</tr>
<tr>
<td>orf00039</td>
<td>24256</td>
<td>26526</td>
<td>+1</td>
<td>13.24</td>
</tr>
<tr>
<td>orf00040</td>
<td>26526</td>
<td>28574</td>
<td>+3</td>
<td>11.64</td>
</tr>
<tr>
<td>orf00043</td>
<td>28574</td>
<td>29440</td>
<td>+2</td>
<td>4.49</td>
</tr>
<tr>
<td>orf00044</td>
<td>29450</td>
<td>29770</td>
<td>+2</td>
<td>11.76</td>
</tr>
<tr>
<td>orf00046</td>
<td>29773</td>
<td>30201</td>
<td>+1</td>
<td>4.90</td>
</tr>
<tr>
<td>orf00048</td>
<td>30210</td>
<td>31193</td>
<td>+3</td>
<td>10.50</td>
</tr>
<tr>
<td>orf00049</td>
<td>31190</td>
<td>31444</td>
<td>+2</td>
<td>7.17</td>
</tr>
<tr>
<td>orf00051</td>
<td>31715</td>
<td>32254</td>
<td>+2</td>
<td>2.39</td>
</tr>
<tr>
<td>orf00053</td>
<td>32238</td>
<td>32693</td>
<td>+3</td>
<td>6.68</td>
</tr>
</tbody>
</table>
Fig. 7.3  Output of a BLASTP search using the BLAST server at NCBI. The amino acid sequence of a predicted phage protein was submitted as a search query to the server, which returned a summary of the search results; a screen shot of the results window is shown. At the top, a diagram shows the correspondence of a putative conserved domain identified in the search with a linear representation of the query sequence. Below that is a diagram of related sequences identified in the search, with the closest matches shown in red. The identities of the specific matches and their scores are shown in the bottom panel (see Color Insert)

- Phage genomes are tightly packed, so call the longest ORF.
- Genes are often found in groups in the forward or reverse directions.
- Genes do not extensively overlap. Overlap by a stop codon is common. However, if the overlap is greater than seven amino acids, then other data will be needed to provide additional support for that start site.
- Two genes are not called in the same space. Unless there is supporting data to the contrary, each segment of DNA only codes for one gene.
- Glimmer choice (start coordinate listed) is valuable.
- GeneMark choice (the beginning of the coding potential) is valuable.
- Shine–Dalgarno score (a feature of DNA Master) – the bigger the score, the better the potential start site.
- Shine–Dalgarno sequence – Translation initiation requires recognition of the correct location by binding of the ribosome to the mRNA. Specific sequences
Fig. 7.4 Representations of genome analyses using DNA Master. (A) The output from the “Frames” routine in DNA Master showing six reading frames, with start (short vertical lines) and stop codons (full height lines) indicated. (B) ORFs with strong prediction for coding potential predicted by GeneMark and Glimmer are designated by a thick green horizontal line. (C) Clicking on a segment of a frame of interest produces a thin green line and also links to the “choose start” routine shown in (D) the ORF and start codon corresponding to the frame segment in (C) is in red. (4). (E) The genome segment highlighted in (C) is revealed as a selected segment (black background) in the “Sequence” panel of DNA Master and can be added to the annotation or copied as either a nucleotide or a translated amino acid sequence for BLAST analyses (see Color Insert).

(Shine–Dalgarno sequence) are recognized and bound by ribosomes. The consensus Shine–Dalgarno sequence in *Escherichia coli* is 5’-AGGAGGA spaced approximately seven bases upstream of the start codon but it is similar for many bacteria. For any given gene both the sequence and the spacing can vary.

- The three possible mycobacterial start codons have an overall distribution of 45% ATG, 45% GTG, and 10% TTG. Glimmer does not identify the TTG starts.
- Blast comparison – copy your call as a translation and submit a BLASTP request to NCBI BLAST. Does your predicted gene match other GenBank entries? If your gene is similar to a known gene but different start locations are chosen, it does not necessarily mean that yours is wrong.
- As the phage genome ORFS change direction, recognize that there probably needs to be intergenic space for promoters.
- Be cautious when identifying very small ORFS. The average length of phage genes is only 200 codons, but there are few smaller than 50 codons.
- Negative frame genes are somewhat more difficult, take time to get oriented.
Fig. 7.5 Six-frame translation of a phage genome segment. DNA Master provides an output showing the DNA sequence of both strands as well as the amino acid sequences in each of the potential six reading frames. The amino acid sequences of the potential ORFs predicted by Glimmer and GeneMark are colored. Each of the potential start codons shown in Fig. 7.4E (and within this DNA segment) are shown boxed (see Color Insert)
• DNA Master stop codons are indicated by full frame vertical lines and start codons as 1/2 the height of the frame.
• Only DNA Master calls TTG starts.
• Check the gene calls at the right end of a circular genome to be sure there is not a gene that wraps around the end of the genome.

These information sources are weighted and evaluated differently in the determination of a gene. This decision-making reasoning process is interpretive (Figs. 7.4 and 7.5). Following is a schematic presenting of this process:

**Required Determination**: Identification of genetic sequences that have a high potential to be genes.

**Sample**: FASTA file – the complete genetic sequence for the bacteriophage genome in a common file format.

**Results and Analysis**: The process of identifying a gene within the complete genetic sequence is done in several steps. The recording of the analysis in its initial stages is conducted using a six-frame translation print out, GeneMark graphic output, and DNA Master. Using all of the programs (Glimmer, GeneMark, DNA Master) and the evaluations listed above, the potential genes of each bacteriophage genome are identified. This process is recorded in the DNA Master file and on the GeneMark and six-frame translation printouts using a marker.

### 7.4 An Educational Assessment Strategy for the Scientific Inquiry Process of Phage Hunting

In order to fulfill the aims of the assessment program and address in a contextualized manner the nature of the scientific inquiry process of bacteriophage isolation and identification, a series of assessment tools utilizing different methods and addressing different knowledge sources were developed. In all, the strategy includes the following five assessment tools: the **Substantive Knowledge Test**, the **Physical Checklist**, the **Visual Literacy Test**, the **Notebook Assessment Tool**, and the **Knowledge Presentation Performance Test**. Taken together this series of assessment tools provides in-depth information on the achievements of the students involved in the scientific inquiry process of phage hunting and also information that can be used to provide students with relevant and timely feedback on their developing scientific inquiries. These tools can be defined as follows:

1. **The Substantive Knowledge Test**: This is a test of the factual knowledge that develops as a result of the specific scientific inquiry experience of working in a microbiology laboratory. This test has the format of a written test designed to elicit factual scientific knowledge. The test can have the format of a multiple-choice or open-ended question format and is parallel to other tests developed within the university courses that relate to substantive knowledge. This test is divided into two sections, one conducted in the middle of the program and one
at the end of the program and is a final summative test of substantive knowledge development. This test involves professional content knowledge relating to bacteriophage. The knowledge assessed by this tool is significant in that it provides information on the factual knowledge of microbiology that has been achieved during this educational scientific inquiry experience.

2. **The Physical Checklist**: Working in a laboratory is a basic aspect of most scientific inquiries. Within the phage-hunting program a significant amount of the student’s work is actually conducted in the microbiology laboratory. Accordingly the development of a tool that assesses basic working knowledge of a laboratory is a desirable aim. This tool addresses the basic physical knowledge that is required to function within the wet laboratory. The tool is designed for the novice workers in the microbiological laboratory and addresses basic techniques that have wide-based application in many of the different protocols used in the stages and steps of the bacteriophage isolation and identification process. An analysis of the physical knowledge utilized in the various stages of the phage-hunting process revealed five physical techniques that were repeated and considered basic to many of the procedures conducted in the microbiology laboratory. The five techniques are Use of a Pipette Gun; Use of a Micropipettor; Centrifuge Use; Aseptic Technique; and Plating Technique. The assessment tool has the format of a checklist specifying the required usage of each of these techniques and differentiating between different aspects of the technique. This assessment tool is used in an on-going manner and in relation to the students’ real work in the phage-hunting program. In other words, the student is not artificially tested; rather, as the student works, the instructor-mentor closely observes the student’s work and “checks-off” on the checklist the quality of the student’s functioning with the different techniques. This happens in relation to every relevant usage of these specific techniques and as such is a rolling procedure in which a series of dated evaluations of physical technique are conducted. The physical checklist is designed as a diagnostic and summative assessment tool. As a diagnostic and formative tool, the checklist provides mentors with specific information on the functioning of the students. If any problems are found, the instructor comments on these and provides instructional feedback in relation to the specific step or stage of the phage-hunting process that the student is conducting. As a summative assessment tool, the checklist shows levels of mastery of these different basic microbiological laboratory techniques and is to be considered the development of procedural scientific knowledge. The knowledge assessed in this tool is significant on three levels: (1) it directly influences the quality of the microbiological research conducted by the student; (2) it provides a basis of physical knowledge concerning microbiological research that allows results to be interpreted and comprehended; and (3) it provides a basis for exploring developments in self-understanding as a researcher.

3. **The Visual Literacy Test**: A central aspect of the microbiological process of phage hunting consists of the ability of the researcher to conduct a visual analysis of various representations used in the microbiology laboratory. To understand and interpret these representations, different knowledge sources need to be
addressed. The interpretation of a visual representation requires the integration of substantive microbiological knowledge and procedural knowledge of the cognitive, representational, and epistemological aspects of microbiological research. Accordingly a test that is based on the interpretation of authentic visual representations used in the phage-hunting process elicits information that can be used to assess the development of the cognitive and representational knowledge that is specific to the phage-hunting process. In addition this test elicits information that can be used to assess the understanding of the microbiology of organisms that are researched and the specific process of scientific inquiry utilized in the phage-hunting program. For the PHIRE program five specific stages were chosen and the visual components of these reasoning processes were analyzed. Five specific visual analysis tasks were chosen for the construction of this test: Phage Isolation: Identification of plaque; Phage Purification: Generating a pure phage population; Phage Amplification: Calculation of lysate concentration (titer), Selection of maximum bacteriophages yield; DNA Comparison: DNA Restriction Enzyme and Electrophoresis: Evaluation of uniqueness of isolated bacteriophage; and Genome Annotation and Analysis: Gene identification. These five tasks cover the length of the phage-hunting process and are directed toward points in the process in which informed decisions need to be made. The format of the visual analysis test consists of a series of specific visual representation and questions designed for each task. The visual analysis test is administered by the student-mentor at the completion of each stage in the phage-hunting process to which it is relevant. As such it comes after the student-researcher has conducted her/his own research and can provide information on the understanding that has been acquired in relation to that stage. Since the visual analysis tasks cover the whole of the phage-hunting program, by considering the complete set of visual analysis tasks a picture of knowledge development in relation to the internal aspects of the phage-hunting program can be constructed. As with the physical checklist, the visual analysis test provides data that can be used for diagnostic and outcomes assessment. The administration of the test by the instructor is an opportunity for an interpretive discussion with the student concerning different ways of understanding the visual representations and can be an opportunity to provide direct feedback and instruction (if necessary) concerning the understanding of the process. The knowledge assessed by this tool is significant on three levels: (1) it provides information that can be used to assess the on-going cognitive development of microbiological understanding concerning the phage hunting program; (2) it provides information on the accumulated cognitive understanding of phage hunting and bacteriophage as an organism; and (3) it provides a basis for exploring developments in self-understanding as a researcher.

4. The Notebook Assessment Tool: The process of conducting a scientific inquiry within a professional research laboratory is a lengthy process that involves a range of physical, cognitive, and representational skills. The role of the laboratory notebook is defined as a scientific report that documents the on-going
scientific activities and thought processes of the researcher. Previous research (Hanauer, 2007) suggests that the notebooks as used by student-researchers may not be fully attuned to the professional understandings of the role of notebooks in the laboratory. The main conclusion from this situation is the need for further clarification with students as to how to use these literacy products as part of the scientific inquiry process. However, as proposed here, the laboratory notebook has the potential as a highly significant assessment tool if appropriate instruction and feedback are provided. As an authentic assessment tool, the scientific notebook indeed provides the instructor with the option of fully following both the educational and the scientific process of phage hunting and also provides an on-going log of all activities and the understanding of these activities. The assessment of the notebook is an on-going activity of the instructor and part of the regular educational process. The knowledge assessed by this tool is significant in that it provides an opportunity to observe the on-going development of the scientific inquiry process and for the instructor to provide timely instruction and feedback on this process.

5. The Knowledge Presentation Performance Test: This test involves the assessment of the student’s ability to synthesize, contextualize, and theorize the results of the scientific inquiry process and present these within an acceptable scientific format. Observation within the laboratory has shown that the most frequent form of scientific inquiry presentation used consists of the conference poster format and accordingly the current assessment proposal suggests that a similar format be used as an assessment tool. As an assessment tool, the presentation of the outcomes of a scientific inquiry process in the form of a conference poster has two main sections: assessment of the conceptualization of the scientific content and assessment of genre form of the poster. As an assessment tool, rubrics for both of these elements have been developed. The knowledge presentation performance test is conducted at the end of the course in a physical format that is very similar to actual conference presentations. The showing of the posters can be public and offers a professional atmosphere that encourages the students’ sense of authenticity. The knowledge assessed by this tool is significant in that it assesses the students’ real understandings of the outcomes of their scientific and educational endeavors.

### 7.5 Weighting and Timing

The assessment strategy presented in this chapter involves the construction of a composite score from five different assessment procedures that are collected at different times during the course of scientific inquiry. The assumption of the authors of this book is that these information sources do not have equal weighting in relation to their relative importance to the educational process. The weighting of
Table 7.1 Assessment Measures, Knowledge Types, Timing, and Weighting of Assessment Measures

<table>
<thead>
<tr>
<th>Assessment procedure</th>
<th>Knowledge assessed</th>
<th>Assessment timing</th>
<th>Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Substantive Knowledge Test</td>
<td>Summative assessment of substantive factual knowledge</td>
<td>Middle and end of program</td>
<td>20%</td>
</tr>
<tr>
<td>The Physical Checklist</td>
<td>Diagnostic, formative, and outcome assessment of physical activities in the wet laboratory</td>
<td>On-going first half of phage-hunting process</td>
<td>5%</td>
</tr>
<tr>
<td>The Visual Literacy Test</td>
<td>Diagnostic, formative, and summative assessment of cognitive and representational aspects of the phage-hunting process</td>
<td>On-going through the program. The five visual analysis tasks cover the length of the phage-hunting processes</td>
<td>25%</td>
</tr>
<tr>
<td>The Notebook Assessment Tool</td>
<td>Diagnostic and formative information relating to substantive, procedural, cognitive, and representational knowledge of the phage-hunting process</td>
<td>On-going throughout the program</td>
<td>25%</td>
</tr>
<tr>
<td>The Knowledge Presentation Performance Test</td>
<td>Summative assessment of presentational knowledge</td>
<td>End of program</td>
<td>25%</td>
</tr>
</tbody>
</table>

Different assessment tools relate to the conceptual understanding of the educational significance of the various components for the scientific inquiry process. Table 7.1 summarizes the roles of the different assessment measures, the timing of the assessment, and the suggested weighting for each item.
Chapter 8
PHIRE Assessment Tools

8.1 Introduction

This chapter provides further clarification and exemplification of the specific assessment tools developed and used within the PHIRE program to evaluate student educational progress. As described in Chap. 7, the overall assessment strategy was based on a contextualized, multimethod approach and consists of five different tools: the substantive knowledge test, the physical checklist, the visual literacy test, the notebook assessment tool, and the knowledge presentation performance test. In the sections that follow each of these assessment tools is described individually.

8.2 The Substantive Knowledge Test

The substantive knowledge test is a tool that assesses the development of factual scientific knowledge as an outcome of the process of scientific inquiry. The substantive knowledge test is the most traditional of the methods used in the current set of assessment tools. The substantive knowledge test used within the PHIRE program consisted of a multiple choice question and open-ended question format and addressed factual knowledge of bacteriophage. As an initial stage, the prime researcher and his team regarded the type of knowledge that they thought could be considered significant as an outcome of the specific scientific inquiry conducted within the PHIRE program. Particular emphasis was placed on factual knowledge that informed the process of scientific inquiry and related to the nature of the organism and the specific processes of scientific inquiry. A list of questions covering the different stages of the PHIRE program and addressing very basic as well as more complex degrees of knowledge was developed. Below are three example questions relating to the early stages of the PHIRE program:

1. What is a bacteriophage?
   a. A bacteriophage is a virus that infects bacterial cells
   b. A bacteriophage is a bacterial cell
c. A bacteriophage is a virus that directly causes disease in humans

d. A bacteriophage is the cause of tuberculosis

2. A bacteriophage can

a. Not metabolize or replicate its own DNA without a host cell
b. Reproduce independently
c. Intake and digest energy from other viruses
d. Needs a food source to survive

3. CaCl₂ is added to media and buffer used for phages and M. smegmatis. Why is Ca²⁺ needed?

a. Ca²⁺ is needed in the metabolic pathways for bacterial growth as well as for phage to bind to cell surfaces
b. Ca²⁺ is needed to stabilize and preserve the solution of phage
c. Ca²⁺ is needed to prevent phages from lysogenizing
d. Ca²⁺ is needed to maintain the structure of phage capsid

8.3 The Physical Checklist

The physical checklist is a tool that assesses basic working knowledge of a laboratory. In the case of the PHIRE program, working with novice undergraduate students with little or no experience of microbiology, the decision was made to focus on core microbiological techniques. This assessment tool is relatively simple to construct as it directly reflects correct usage of equipment within the setting of the professional laboratory and in many cases this has already been described in a protocol handbook used within the laboratory. The idea behind the use of this assessment tool is also quite simple. Rather than setting up an artificial context for tool usage, the idea is that actual work conducted by the student-researcher during the process of scientific inquiry is observed by the scientist-educator and assessed. The physical checklist is an aid in allowing the scientist-educator to assess and record in a systematic way the physical, procedural knowledge of the student-researcher. In this sense the physical checklist is a performance test that is conducted as the students actually conduct their scientific inquiry.

As an assessment tool that is practical to use, several quite simple procedures need to be followed in order to transform the existing knowledge concerning physical work within the laboratory into an assessment tool. The first simple action that needs to be taken is to decide on the specific physical laboratory actions that are important to assess as part of the scientific inquiry process. Naturally this will change in relation to each scientific inquiry process and level of student-researcher. Once this list of actions, procedures, and equipment that needs to be addressed from an assessment perspective has been generated, the second and most important transformation is to delineate, in a checklist format, what correct usage of a piece of
equipment or a specific procedure actually consists of. The detailed explanation of
the checklist, set out as a user-friendly table, is designed to provide the assessor with
a basic set of ideas of what is important to consider in relation to the student’s phys-
ical work in the laboratory and is ultimately the instrument that is used to record the
students progress and understanding of the physical environment and actions within
the laboratory. Once the checklist has been developed decisions need to be made as
to when to assess each of the procedures described on the checklist. This decision is
directed through the empirical analysis of the physical knowledge required for each
of the representational milestone stages of the scientific inquiry process. To most
prime researchers, this will be an obvious decision.

To exemplify the nature and form of the checklist, below is an example of the
checklist developed for the basic microbiological technique of plating.

<table>
<thead>
<tr>
<th>Physical technique</th>
<th>Subcategory</th>
<th>Correct usage</th>
<th>Perceived problems</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plating technique</td>
<td>Overall usage</td>
<td>Holds pipette gun at 45 degree angle</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control bubbles when dispensing media</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spreads liquid evenly</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct labeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Allows plates to cool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incubates plates according to protocol</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uses warm top agar</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Appropriate use of micropipette or pipette gun</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8.4 The Visual Literacy Test

The visual literacy test is a tool that is designed to assess the contextualized and
integrated understanding of the specific scientific inquiry process that is at the heart
of the educational program. As set out in the first sections of this book, the pro-
cess of scientific inquiry can be set out as a series of representations with associated
cognitive knowledge and physical actions. The aim of the visual literacy test is to
use these milestone representations as significant points of assessment and to con-
sider what it is that the student actually understands about this scientific process.
This method of assessing scientific inquiry knowledge is designed to differentiate
between those students who have good control over the mechanical components of
the laboratory (as assessed in the physical checklist tool presented above) and the
actual conceptual understanding of the process of scientific inquiry. Accordingly,
this test integrates a visual analysis of representations in the microbiology labora-
tory with both cognitive and representational knowledge.

This test, more than any of the tools described in this chapter, is directly based
on the empirical description of the specific scientific inquiry process that is used in
the educational program. In the last chapter, it was specified that five representa-
tional milestones were identified for the PHIRE program: Step 1 – Phage Isolation:
Identification of plaques; Step 2 – Phage Purification: Generating a pure phage population; Step 3 – Phage Amplification and Purification: Calculation of lysate concentration: Selection of maximum bacteriophages yield; Step 5 – DNA Comparison: DNA Restriction Enzyme and Electrophoresis: Comparison of electrophoresis gel and decision as to value of continuation of process; Step 7 – Genome Annotation and Analysis: Gene Calling – Selection and identification of genes. A specific test was designed for each of these stages with the assumption that student-researchers would be tested following the completion of the in-lab work. A detailed analysis of the cognitive and representational aspects of each of these stages was conducted. In order to aid in the development of an integrated assessment tool, a basic knowledge and cognitive abilities template was designed based on the analysis of the specific representations. For the four wet-lab tasks, this template consists of

1. Knowledge of an organism’s characteristics
2. Understanding of the manipulated environment
3. An ability to selectively focus on specific visual information in the representation
4. The ability to infer the state of the organism from the visual analysis through an argument structure
5. The ability to make an informed decision on future action that is required

This basic knowledge and set of abilities provide the structure of information that needs to be assessed in relation to these tasks and was reflected in the questions for the visual analysis task.

The last task – Gene Identification – is different from the first four tasks in that its decision-making structure uses symbolic core characteristics and is an interpretive, knowledge integration task. This template specifies basic knowledge and abilities that are required for making a decision in this setting:

1. Knowledge of the characteristics of an organism
2. Knowledge of the translation of the organism’s characteristics into symbolic form
3. Understanding of the workings and outputs of the different computer programs
4. Understanding of the nature of the basic analysis conducted by the different computer programs which led to the specific output that is produced
5. The ability to weigh and integrate different information sources in making an informed decision.

This basic knowledge and set of abilities provide the structure of information that needs to be assessed in relation to the gene identification task.

The form of the visual literacy test is an open-ended question format dealing with the visual analysis of presented laboratory representations in the form of digital photographs. The representations were drawn from work actually conducted in the laboratory and the questions were designed according to the question template above. The result is a test which includes a visual representation and a set of associated questions. For example consider the following page from the test of the Phage Isolation stage (Fig. 8.1).
Visual Literacy Test One: Identification of Bacteriophage

Look carefully at the picture of Figure 8.1 and answer the four questions listed below. Questions:

1. Are there any bacteriophages on this plate?
2. How do you know if a bacteriophage is present on a plate? Specify exactly what evidence from the plate you are using to make this decision.
3. The plate in the picture is covered with bacterial cells. Why is this important and how does it help you to identify the presence or absence of bacteriophages?
4. What is the next stage of working with this specific plate?

As can be seen in this example, a digital photograph of an authentic outcome has been presented and is the basis for the set of questions that is being asked. The questions address the understanding of the representation and are designed around the analysis of the cognitive and factual knowledge required to actually understand and conduct this specific scientific inquiry. Naturally, a specific response rubric was developed for each of the visual analysis tests following the piloting of the materials.

8.5 The Notebook Assessment Tool

Scientific notebooks have enormous potential as educational and assessment tools of the process of scientific inquiry. This potential is based on two central ideas: (1) the scientific notebook is an authentic part of professional scientific inquiry; (2) the scientific notebook provides an in-depth, up-to-date description of the procedures, outcomes, and thoughts concerning the process of scientific inquiry conducted by the student (Hargrove & Nesbit, 2003). The notebook assessment tool is designed to
collect data on the whole of the scientific inquiry process conducted by the student. Specifically, the aims of the laboratory notebook assessment procedure are

1. To elicit assessment data that can be used to determine the levels of procedural and representational knowledge development achieved by the participants in this program.
2. To provide diagnostic information that can be used to enhance each student’s understanding of the substantive, procedural, and representational aspects of the bacteriophage hunting process and formative information to the scientist-educator so that changes can be made to the educational plan if necessary.
3. To provide summative data on the specific aspects of the scientific microbiological inquiry process of the isolation and characterization of bacteriophage

It is should be noted that the use of laboratory notebooks as part of the educational assessment process does not guarantee the collection of useful summative, diagnostic, and formative data. Hanauer (2007) and Ruiz-Primo, Li, and Shavelson (2002) note that within the notebooks they studied, the recording of the procedures and results were problematic leading to a situation of difficulties in replicating or directly following what the students were doing or how they understood their studies. In addition, in many cases the results were not written in a way that would allow conclusions based on evidence or patterns of significance to emerge. The notebooks did not reach their full scientific and educational potential as a full record of the scientific and thinking aspects of scientific work.

One of the central issues in utilizing laboratory notebooks as assessment tools concerns the role of the instructor. The instructor’s role in properly introducing the laboratory notebook as a scientific literacy practice and providing on-going feedback on the quality of the science and the description of the entries is pivotal in making the laboratory notebook a valuable scientific and educational tool. The role of the instructor can be divided into three time periods: (1) at the very beginning of the program, the introduction of laboratory notebook writing; (2) throughout the program, the provision of feedback in relation to the quality of the science and notebook entries in describing the scientific process; and (3) at the end of the program, the assessment of the completed laboratory notebook. The full protocol for the usage of the notebook as a scientific and educational tool is provided below in relation to these three time periods.

### 8.5.1 Program Beginnings and Notebook Introduction

At the very beginning of the scientific inquiry program the presence and usage of laboratory notebooks needs to be introduced. This introduction is crucial in that it defines the aims and usage of this specific literacy practice. In our assessment procedure we used the following handout for this purpose.
Principles of Notebook Writing

1. **The Principle of Autonomous Replication:** The main scientific principle directing the writing of laboratory notebooks is that the scientific inquiry that is described in the notebook can be followed and understood autonomously by an outside reader without the need to search for additional information from other sources. In other words, *the notebook stands as an autonomous, comprehensive, and explicit description of the scientific inquiry that was conducted and the results that were achieved.* A basic test of the quality of the notebook is that, if necessary, the outside reader could replicate the whole study and reach the same results.

2. **Full Presentation of the Scientific Argument:** In order for the outside reader to be able to fully follow the specific details of the scientific inquiry process, the entries made in the notebook need to construct a full scientific argument in relation to each set of specific procedures and results that are presented. The scientific argument consists of the following entries that need to be included and presented in this specific order:

- **Aims and Purposes of the Proposed Procedure:** Before presenting any specific scientific procedure, *the aims of the proposed methodology need to be explicitly stated in the notebook.* This might be as simple a statement as “I need to clean my samples from excess dirt so that I can later be able to plate them.” Or in some cases the description of the aim of the procedure may include the specification of a quite complex problem that needs an experimental design to differentiate between different options. In other words, the student-researcher needs to explicitly answer the question as to why any specific procedure is being conducted. The presentation of the reasons why a procedure was conducted allows an autonomous outsider reader to follow the thoughts of the student-researcher in conducting this procedure and the decision-making processes that were involved.

- **The Accurate and Comprehensive Reporting of a Scientific Procedure:** A basic aspect of any notebook is the full, explicit and accurate reporting of the procedures that were conducted. In principle a full description of every action taken in relation to the scientific procedure is required. In some cases the student may be following a well-known protocol and every step is followed in exactly the way it is described in the protocol. In this case reference can be made to the protocol. It is important to *remember that an outside reader must be able to follow exactly what was done in the scientific inquiry without the presence of the writer and just by reading the description of the procedure followed.* In some cases the description of the scientific procedure followed may include visual elements such as hand-drawn pictures. Whether the description is visual or verbal it is important that it is comprehensive, accurate, and explicit. The student-researcher needs to answer the question – How was the scientific inquiry completed? An explicit answer to this question provides a full description of the procedure.
• **Accurate Reporting of Results:** The accurate reporting of the results found within the scientific inquiry process and as an outcome of the specific procedures described in the notebook is a core aspect of the laboratory notebook. *The description of the results needs to be as accurate and as comprehensive as possible.* In many cases within the wet laboratory, this may include digital photographs or other visual representations of the actual results themselves. All visual material must be properly labeled and of a high quality. In some cases you may need to provide a mathematical description of the results that a specific procedure or experiment achieved. These results need to be detailed and present a full picture of what was found. The basic principle here is that the outside reader can fully comprehend what the actual results of the procedure were. With this entry, the student-researcher needs to answer the question – *what was the outcome of the application of this specific procedure?* An explicit answer to this question will provide a full description of the results that were found.

• **Interpreting Data and Reaching Conclusions:** Having presented an accurate and full description of the results, it is necessary to explain what the results mean. The results need to be interpreted and conclusions for the continuation of the scientific inquiry process discussed. The interpretation of the results leads directly to the specification of an aim for the next stage of the scientific inquiry process. It is particularly important to report on the presence of any unexpected findings and problems that arise. The interpretation presents the student-researchers understanding of what the results show and what this tells her/him about the scientific inquiry process they are involved with. Basically the student is trying to answer two different types of question: *What do the results mean?* and *What do I need to do now?* The first question relates to the results themselves and the second question connects to the continuation of the scientific inquiry process. Table 8.2 is a schematic representation of the types of notebook entry that constitute the scientific argument and the relationships among them.

### 8.5.2 On-Going Laboratory Notebook Usage and the Provision of Feedback

Good laboratory notebook writing practices will develop if the scientist-educator reads the notebooks and actively provides feedback to the student-researcher on the quality of the notebook from a scientific and educational perspective. This is particularly important in the first third of the scientific inquiry process as this is the time at which long-term scientific writing behaviors can be encouraged and normalized. During this initial period, the notebook is reviewed at on a weekly (perhaps bi-weekly) basis and explicit face-to-face instruction provided to the student. In order to facilitate this process we developed the rubric for laboratory notebook assessment presented below (Tables 8.1 and 8.2):
Table 8.1  Rubric for assessment of laboratory notebooks (scientific argument)

<table>
<thead>
<tr>
<th>Notebook entry categories</th>
<th>Inadequate</th>
<th>Adequate</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement of purpose of procedure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Description of procedure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal presentation of results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual presentation of results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpretation of results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presentation of thoughts, problems, and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>deliberations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall evaluation of reporting of</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>experimental procedure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As can be seen the rubric relates to the elements of the scientific argument that are present within the laboratory notebook. Particular attention is directed at the full contextualization of the procedures and results that are presented. Previous research has shown that students do not usually enter these into their notebooks.

8.5.3  Final Assessment of Laboratory Notebooks

*(End of Program)*

As weighted within the current assessment strategy, the laboratory notebook constitutes 25% of the final assessment of the student. As a component in the final assessment the notebook needs to be assessed for its value as a scientific document and for its ability to provide the instructor with information on the student’s demonstrated knowledge of substantive, procedural, and representational aspects of the process of bacteriophage isolation and genomic annotation. At the end of the program, the laboratory notebook is used as a scientific and educational document that allows the assessment of the outcomes of the scientific inquiry process that the student-researcher was involved with. To simplify this educational assessment process a series of specific questions that the researcher-instructor can ask following the reading of the laboratory notebook were developed. These questions are designed to direct the instructor’s attention to specific issues in the assessment of the notebooks. The final assessment of the laboratory notebook is based on the instructors answer to the following questions:

1. Does the notebook provide you with sufficient information that would allow you to replicate the study if necessary? Does it allow you to fully understand the results that are presented and the specific procedures that led to these outcomes? *(Validity as a scientific document)*

2. Does the notebook provide you with sufficient information that allows you to follow and clearly understand all the different stages of the presented scientific inquiry? *(Validity as a scientific document)*
<table>
<thead>
<tr>
<th>Notebook entry categories</th>
<th>Inadequate</th>
<th>Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement of purpose of procedure</td>
<td>Does not provide or provides in a limited fashion a statement of the purpose and aims of the procedure that is to be followed</td>
<td>Provides a clear reason and explanation of the procedure to be followed. The reason given is accurate within the context of the type of scientific inquiry that is being conducted</td>
</tr>
<tr>
<td>Description of procedure</td>
<td>The description of the procedure does not allow the reader to understand exactly what was done by the researcher. The description of the procedure is very limited and partial</td>
<td>The description provides a clear and detailed description of the procedure followed so that the reader can understand and be able to replicate exactly what was done by the researcher</td>
</tr>
<tr>
<td>Verbal presentation of results</td>
<td>The reader has difficulty in understanding what the results are and how they relate to the procedure. This may result from a very limited or very partial presentation of results or the lack of clear connection between results and procedures described</td>
<td>The verbal presentation of results provides information that allows the reader to fully comprehend the results that were found. In some cases the verbal results may be an annotation attached to a visual presentation. But taken together the reader understands exactly what results were achieved from the conducted procedure</td>
</tr>
<tr>
<td>Visual presentation of results</td>
<td>The photographs, pictures, drawings, or graphs: are not of a quality that allows results to be understood; are not properly labeled; only partial results have been attached</td>
<td>The photographs, pictures, drawings, or graphs are presented in manner that allows the viewer to fully comprehend the results that were achieved. The visuals are properly labeled and presented in a visually clear and coherent manner. The visuals are of a good quality that allows a detailed examination</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Table 8.2 (continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interpretation of results</strong></td>
</tr>
<tr>
<td>There is a very limited (if any) presentation of the researchers’ interpretation of the results and what new steps need to be taken to continue the process of scientific inquiry.</td>
</tr>
<tr>
<td>The researcher writes out her/his interpretation of what the results mean and how they reflect the scientific procedure followed. In addition, the researcher explicates what new steps need to be taken following this interpretation of results.</td>
</tr>
<tr>
<td><strong>Presentation of thoughts, problems, and deliberations</strong></td>
</tr>
<tr>
<td>There are very few (if any) comments relating to the research thought processes, problems, or deliberations. The researcher may be using the notebook to make personal comments which are not relevant to the scientific inquiry process itself.</td>
</tr>
<tr>
<td>Throughout the description of the scientific inquiry process, the researcher explicates in written form the problems, deliberations, and thoughts that arise in relation to the scientific process itself. This series of entries allows the reader of the notebook to follow the thought processes of the researcher while conducting this scientific inquiry. These comments may include emotive presentation dealing with the results of procedures and the excitement of discovery or frustration of the unexpected.</td>
</tr>
<tr>
<td><strong>Overall evaluation of reporting of experimental procedure</strong></td>
</tr>
<tr>
<td>Following the reading of the notebook entries the reader was not able to reconstruct the exact procedures, results, understandings, and thought processes used by the researcher in conducting their scientific inquiry. The reader was involved in an extensive process of inference and guessing in order to fill in the missing information. Signs of confusion as to what scientific process was conducted, what was found, or what it meant to the researcher are signs that the notebook description is lacking.</td>
</tr>
<tr>
<td>Following the reading of the notebook entries the reader was able to fully reconstruct the whole process of scientific inquiry including the reasons for the procedure, the exact procedures that were followed, the details of the results, and the researcher’s understandings of what was found. The scientific study described in the notebook is fully replicable from the description that is provided.</td>
</tr>
</tbody>
</table>
3. Based on the evidence presented in this notebook, is it your evaluation that the student-researcher has a good understanding of the procedural and representational aspects of the bacteriophage isolation and genomic annotation? (*validity as a scientific document*)

### 8.6 The Knowledge Presentation Performance Test

The knowledge presentation performance test involves an authentic summative assessment of the scientific inquiry process that the student-researcher conducted. The knowledge presentation performance test is basically the public presentation of a conference poster constructed and designed by the student. Discussion with the scientist-educator is considered a central part of the presentation process and the point at which the assessment data are collected. A professionally produced conference poster presents information in a coherent, concise, and relevant manner and allows the instructor to assess the student’s ability to synthesize, contextualize, and theorize the results of the scientific inquiry process and present these within an acceptable scientific format. The aims of the knowledge presentation performance test are:

1. To elicit assessment data that can be used to determine the scientific outcomes and value of the student’s research
2. To elicit summative assessment data concerning the student’s understanding of the substantive scientific aspects of their scientific inquiry process
3. To elicit assessment data on the student’s ability to synthesize, contextualize, and theorize the results of their scientific inquiry process
4. To elicit assessment data on the student’s ability to present their research in accordance with the guidelines of professional scientific presentation

Conference posters have enormous potential as both educational and assessment tools. The advantage of the conference poster is that this is a real genre that students can actually produce and use to summarize their research. However, for this to work within the context of an educational program and as an assessment tool, the researcher-educator has a crucial role to play in creating an authentic context for the presentation and discussion of research. As with any research project, the conference poster comes only after significant data have been collected and is a form of conceptualization and summary of findings and outcomes. Accordingly, the knowledge presentation performance test is used at the very end of the undergraduate scientific inquiry process. Discussion of the conference poster is conducted during the last quarter of the scientific inquiry process. Basically there are two stages to be considered in the usage of the conference poster as an assessment tool: the development stage and the presentation stage. The full protocol for the usage of the conference poster as a scientific and educational tool is provided below in relation to these two stages.
8.6 The Knowledge Presentation Performance Test

8.6.1 Conference Poster Development

The development of a conference poster is a reflective and summative process that transforms the raw material of the scientific inquiry described in the laboratory notebook and the associated visual representations into a focused presentation of a significant scientific outcome. This process of transformation is highly valuable from an educational perspective in that this is one possible point at which the student can really consider what she/he has found out and relate this to understandings within the wider scientific discipline. The process of developing a conference poster incorporates this educational aspect of personal reflection on the scientific inquiry process. Hanauer (2007) defines five stages of poster development: development of poster genre knowledge; reflection on notebook and visual data; production of a written theme or core discovery of the inquiry process; construction of a conference poster; and construction of an accompanying oral explanation. In our assessment process we used these stages to design the educational process of conference poster development. These stages are as follows:

1. Development of Conference Poster Genre Knowledge: As a first stage of knowledge development the instructor presented the ASM (American Society of Microbiology) criteria for conference poster presentations. The initial introduction specified the elements that needed to be in the poster including an abstract, background, methods, results, conclusions, future directions, and references. This was followed by the analysis of real conference posters from professional meetings.

2. Reflection and Discussion of Notebook and Visual Data: Time was allotted for the student to carefully consider her/his notebooks and associated scientific inquiry data. Of particular importance was the consideration of the different visual representations that she/he has collected as part of her/his inquiry process. The student was directed to carefully consider all of her/his data and decide what she/he thought is important to present at a professional conference. At this stage, consultation with the scientist-educator and other peers could be valuable in helping the student to understand the value of her or his research project.

3. Production of a Written Narrative: Following discussion with the instructor students were required to produce a brief abstract of the main directions of the conference poster. The students were instructed that the brief abstract presents the main findings of the scientific inquiry that the student conducted. The student chose those visual representations which were most relevant in the presentation of these specific findings. It is important that the student fully contextualize the findings with the framework of a scientific argument that provides the reader with access to the background, methods, results, and conclusions concerning the study. This written abstract is presented to the instructor and subsequently discussed.

4. Development of Conference Poster: The process of actually developing the conference poster involves making good design decisions based on the abstract that has been written and discussed with the instructor. It is important to specify and
reinforce the sections that need to be present with the conference poster and that the visual representations have to be carefully chosen to provide the most relevant visuals for the poster format. Basically the conference needs to have a title, the name of the researcher, an abstract, background, method, results, conclusions, future research directions, and references. In addition the results sections need to present the appropriate visual representations for the specific findings to be presented. The student needs to consider the appropriate organization of the poster so that it flows correctly and provides the viewer with all required information. The use of poster computer programs such as Canvas is to be encouraged as it allows cut and paste options and the ability to try out different presentation styles. It is important to specify that the poster needs to be professionally presented.

5. Development of the Accompanying Oral Statement: Once the poster has been printed out the student was instructed to develop a succinct 5 min description and explanation of the poster. The description includes a statement about the main findings of the inquiry, the methods used to reach these findings, and the significance of the study. A good way of conceptualizing this process for the student is to think of the oral presentation as “taking your viewer on a guided tour of the poster and its associated scientific inquiry process.” The oral presentation is written out and practiced by the student so that it is fluent.

8.6.2 Conference Poster Presentation

In order for the conference poster to function as an authentic scientific process, the instructor needs to construct a setting in which the conference posters are presented. It would be best if this was not during class time and included other members of the faculty and student bodies. The conference poster session can be simply the designation of a room and time for the presentation of the conference posters. This can be publicized so that there is access to this research by a variety of university members. Once a setting has been found students are told that the knowledge presentation performance test will be conducted during this open poster presentation session. The assessment of the conference posters consists of three components: assessment of the scientific inquiry; assessment of student understandings; and assessment of presentation abilities.

Assessment of Scientific Inquiry: In order to assess the value of the scientific inquiry process conducted by the student-researcher, the instructor considers whether the information collected and presented by the student has real scientific value within the scientific discipline. On a very basic level, the value of a scientific inquiry can be that it provides new data, new methods, or new insights. For example, it is assumed that the contribution of the majority of the student posters dealing with phage hunting would be on the level of new data concerning unique bacteriophages. The question the assessor needs to ask is

1. Does the data presented in the student conference poster really involve a development in scientific knowledge?
This question is at the core of all scientific inquiry and accordingly is significant in the evaluation of the final product of this specific scientific inquiry process – the conference poster.

Assessment of Students’ Understandings: As part of the knowledge presentation performance test the student will be required to develop a poster, an oral statement concerning the poster, and provide interactive responses to instructor questions concerning her/his presentation. Having carefully viewed the poster, heard the oral presentation, and interacted with the student concerning the poster, the assessor considers the following questions:

2. Do the conference poster and the oral presentation fully contextualize the scientific findings that they present?
3. Having viewed the conference poster and heard the oral presentation are you convinced that the student fully understands the results that they have presented and the specific methodological procedures that were used to produce these results?
4. Does the student understand the importance and scientific value of the findings?
5. Having viewed the conference poster and having heard the oral presentation are you convinced that the students have a good understanding of the scientific inquiry process that is presented on the poster.

Assessment of Presentation Abilities: In addition to the scientific value of the inquiry and the student’s scientific understanding of their specific inquiry process, the assessor can assess the quality of the conference poster and oral presentation. The quality of the poster addresses both the esthetic and informative aspects of the poster and oral presentation. The instructor can ask herself/himself the following questions:

6. Does the conference poster present information in a coherent, concise, and relevant manner?
7. Is the accompanying oral presentation coherent, concise, and informative?
8. Do the poster and the accompanying oral presentation synthesize, contextualize, and theorize the results of the scientific inquiry process and present these in formats that are acceptable to a scientific audience?
9. Is the poster esthetically pleasing and professionally produced?
10. Is the oral presentation fluent and articulate?

The final assessment of the conference poster takes into account all the information addressed in the 10 questions presented above. This is a composite score that addresses the quality of the scientific inquiry, the students understanding of this process, and the ability of the student to present this information in a format relevant to professional science.
Chapter 9
Reflections from an Active Scientist-Educator

9.1 Introduction

In the previous three chapters we have presented and discussed the PHIRE program as a case study for the development of strategies and tools for assessment. In this final chapter, one of us (GFH) will share his perspectives from the vantage point as head of a research laboratory that developed the PHIRE program and – together with my co-authors – developed the assessment procedures associated with it. A basic assumption of all our educational work is that real science is and should be at the heart of any science education program. For us science education and the actual development of valuable scientific knowledge should be integrated. The premise I will discuss in this chapter is straightforward: that there are a multitude of exciting opportunities to develop structured educational programs taking advantage of the research environment and incorporating active assessment procedures. The creation of new science-based, education programs would move both science and education forward. However, many of these opportunities remain unfilled, and while there may be many reasons for this, the barrier doesn’t need to be nearly as high as it appears to be. My experience with the PHIRE program and the assessment tools we have developed offer some useful guidelines that I hope will be helpful to other scientist-educators as they contemplate developing their own programs.

Although the PHIRE program has many attributes that we like – and are described in Chap. 6 – it is certainly not unique. It is a component within a broader effort to fully embrace scientific inquiry with our educational missions, and there have been many other successful efforts toward this goal. Science education in the United States is in a peculiar position; on the one hand, K-12 science education is reported as underachieving greatly, but on the other hand, the United States continues to dominate the world of scientific discovery and publication in the peer-reviewed scientific literature. How can this be? While the roots of this paradox may be complex, programs such as PHIRE have the potential to change the game, extending positive advances to both the research and educational enterprises.

I have presented and discussed the PHIRE program in several different forums over the past couple of years and some common questions often arise. How do you find the space? How do you find the time? How do you pay for this? How do you
select students for the PHIRE program? Doesn’t this detract from your research productivity? What does your institution think of this? Will the PHIRE program work for me? How are the students and program assessed? I will try to address these and other questions here that may help to lower the barriers for other scientist-educators to explore exciting new programs and projects in science and education in their laboratories and classrooms.

9.2 Two for the Price of One: Integrating the Missions

I do not consider PHIRE an education program. At first glance, a program that provides laboratory opportunities to high school students might seem to naturally fit into that classification. But it could be better described as a research program that is particularly well-tailored as an introduction to scientific research. This “introduction” is aimed at undergraduate students, high school teachers, and other novice scientists (journalists, public officials, etc.), in addition to high school students. Perhaps such programs are not appropriate for all life science laboratories, but I’m sure there are many that are. Furthermore, within many research laboratories – including my own – only a subset of the experimental approaches may be suited to this introductory category. The critical issue though is that the project should be commensurate with the on-going research directions and not at odds with them. This is good news for scientist-educators thinking about establishing such programs, since it is likely to be a great deal easier initiating a project situated within the strength of their research expertise, rather than trying to start something new solely from the perspective of educational advantage. The context of the research laboratory, and the professional research conducted therein, is therefore of paramount importance, a theme that penetrates much of the book.

An excellent starting point for the development of an educational scientific inquiry program and its associated active assessment components is for the scientist/educator to conduct a brief self-review of the short-, medium-, and long-term research goals and professional scientific activities. It may also be helpful to review the population of students that you might target and the space available in – or close to – the research lab. Having done this, a series of potential projects could be considered, evaluating them for their suitability as an introductory nature – i.e., those that do not require a lot of prior scientific knowledge, prior research experience, or technical skill. In my experience, these projects are very different from the types of projects that most principle investigators are used to formulating for postdoctoral researchers and graduate students. They are also different to the types of projects that are often suggested to individual independent undergraduate researchers, whose projects are often scaled down versions that might be proposed to a beginning graduate student. Programs such as PHIRE aim to engage 8–12 students at any one time and suitable projects are likely to encompass many of the project attributes discussed in Chap. 6, such as project ownership, parallel project structures, peer mentoring, and progression from concrete beginnings to abstract processing.

If scientist-educators can devise a project that meets their research goals, many of the other essential elements, including active assessment, may well fall into
place. Coming up with the right project is the key then to integrating the educational and research missions, with good prospects for advancing both missions. An interesting aspect is that once a project is established then the assessment goals and practices become much clearer. As described in this book, the process of active assessment is tied directly to an analysis of the knowledge, visual representations, laboratory procedures, and conceptual understandings that comprise the specific types of scientific inquiry utilized within the research laboratory and offer a contextualized approach to assessment. For us, both the educational and assessment aspects of the scientific inquiry program should evolve from the core scientific activity of the prime researcher and the laboratory setting. It is also important to note that as happened in our case, the educational projects can take on a life of their own, presenting new research directions that either the scientist-educator or a collaborator specifically interested in educational and learning issues can advance. Thus, a new line of research emerges from within the laboratory setting, one that ultimately enhances interest in science and the specific research agendas of particular laboratories.

9.3 Some Practical Concerns

One of our aims of this book is to promote the usage of in-laboratory science education programs. From my perspective as a scientist and scientist-educator, I know that the practical issue of setting up a program of this type may often seem formidable and there are indeed numerous practical issues that arise when considering the establishment of research-educational programs such as PHIRE. To ease this thinking process, I would like to share some of my thoughts on the practical issues of space, time, and money.

9.3.1 Space: The Final Frontier

Identifying appropriate laboratory space can be one of the toughest of these issues to satisfactorily address. I have been fortunate to have access to a small laboratory that is adjacent to the rest of my lab space and is physically adjoined to it. This has turned out to be ideal, since it is close enough that the PHIRE students are integrated into many normal operations of the laboratory and the use of common equipment, are sufficiently displaced so as not to “take over” the other functions of the laboratory, and work as a cohort in which they can readily work together and interact. In general, the amount of space is perhaps less critical than its location relative to other research activities. One bench at the end of a laboratory or a small lab displaced further away might also work, and given that space is often a prized commodity the investigator has to get whatever she/he can acquire. In general, when it comes to requesting space from a higher authority in the institution such as a Dean or a Department Chair, the scientist-educator has a distinct advantage, since the educational missions are often more closely aligned with the broader goals of the
institution, and undergraduate research has become a strong selling point as universities and colleges compete for talented undergraduate students. Thus requesting space for an organized undergraduate research program may well be more effective than asking for more pure research space, which typically results in a discussion of how many research dollars are being brought in to support the research.

### 9.3.2 Money and Time

Running a program such as PHIRE is certainly not free and the shape and nature of the program are likely to depend on what financial support can be mustered. There are two main components. Perhaps the most critical is the need for funds to employ a program coordinator who can provide the direction and responsibility for overall implementation of the program. The coordinator is critical, since she/he can assist in choosing students for the program, organizing and mediating group meetings, addressing research problems, and all aspects of coordination. There are many nitty-gritty issues in implementation that the coordinator can and should be responsible for, and if all of these fall to the investigator, then the program may be difficult to sustain. This is central to the issue of time management of the researcher-educator, as running a program like PHIRE could become so time-consuming that it detracts from other duties and responsibilities. Obviously, the specific arrangements will depend on the size of the program and the number of students, but assistance with program coordination is not negotiable. The good news is that with a fully integrated education-research program there is a broad range of sources that can be explored for financial support, including research grants, local foundations, education-based federal grants, and the home institution. This is another example in which the educational components advance the research goals and vice versa.

The second main funding issue is that of student stipends. This is primarily an issue when students are working in the summer term, since most of them cannot afford to forego the opportunity to earn the money that they need to continue their education. Depending on the number of students, this can of course become quite costly, although once again, there can be quite substantial research advances in this time period, provided that a suitable project is in place. The same sources of funds can be explored as described above. If the program runs primarily through the academic term and if external funding is not available, then students could enroll in such a program for credit toward their degree.

### 9.4 Why Bother with Assessment?

A well-formulated project that is well suited to inexperienced students can be expected to have considerable educational advantages, and in some senses the more inexperienced the students are, the greater the education benefits that can accrue. But how do you know if any educational benefits are occurring at all and what is happening to students’ understanding of the project and science in general beyond
9.5 Program Attributes Revisited

the simple accomplishment of assigned technical manipulations. Most scientist-educators are well positioned to assess whether the research goals are being met, but perhaps less so than when it comes to determining whether the educational goals are being achieved. Broadly and as set out in the principles of scientific teaching (Handelsman, Miller, & Pfund, 2007) the educational elements should be considered similarly to research elements, as experiments that are conducted whose results need to be evaluated. Active assessment, as described in this book, provides a conceptual framework and a method for collecting data that can be used to directly assess the value of the program in relation to educational outcomes. Rather than an external and abstracted approach to the assessment of knowledge, in this book we propose high emphasis on authentic, contextualized assessment directed by the prime researcher’s extensive understanding of the scientific inquiry that is at the heart of the educational program. This approach to assessment should allow the prime researcher to collect data that allow both the assessment of the educational outcomes and controlling the actual quality of the research that is conducted in the laboratory, an issue that is always of prime concern for any scientist working with undergraduate researchers. On a different level, any information obtained that demonstrates the effectiveness of the education (or reveals program aspects that function less well) can not only be used to enhance and enrich the program, but also present a compelling argument when requesting support for program continuation.

As described throughout this book, assessment of research-education projects does not have to be onerous provided that the assessment process is situated within the context of both the educational and research missions. I have found it to be an asset to my understanding of what my student-researchers are actually doing and learning in my laboratory.

There has been a substantial amount of research aimed at assessing the educational value of research experiences, and there is broad and general agreement that research is educationally valuable for undergraduate students (Handelsman, Miller, & Pfund, 2007). In general, the question as to whether undergraduate research is educationally advantageous seems to be well-resolved in the affirmative (Lopatto, 2004; 2007). Rather than attempting to revisit this large and important question, the key areas of assessment for the scientist-educator are those directly related to the context in which your specific program is conducted. In other words, consider assessment in the terms developed within this book as a way of really understanding undergraduate research conducted within your laboratory and as part of the process of the development of scientific knowledge. The previous two chapters give a clear description of how we have developed and implemented active assessment tools in the context of the PHIRE program.

9.5 Program Attributes Revisited

In order to further enhance the process of developing research-based scientific inquiry programs, I would like to revisit some of the attributes that I think are central to this type of educational program. In Chap. 6 we discussed a set of attributes
that we had identified in the PHIRE program that we propose as guidelines in the development of new projects. One of these is the idea of project ownership, an element that is of importance since it fuels the self-motivation and independence that facilitate full engagement of students in their projects. A notable example is the program developed by Scott Strobel at Yale University in which students gain project ownership through the isolation and characterization of novel endophytes with interesting and unexpected properties (Strobel & Strobel 2007). A second and related key attribute is that students should be performing authentic academic research. On the whole, students can readily distinguish a laboratory exercise from an authentic research investigation, and the latter has considerable impact on activating the needed self-motivation. An excellent example is the program established by Utpal Banerjee at UCLA in which a substantial group of undergraduate students authored a peer-reviewed publication (Call et al., 2007; Chen et al., 2005). Lastly, parallel project structures offer tremendous attributes in that the commonality of techniques and the inherent compatibility with peer- and near peer-mentoring systems build coherence within the project, provide a common ground for students to discuss their projects, and facilitate a variety of assessment strategies.

9.6 Summary and Conclusions

My involvement in the PHIRE program has been a real joy. Observing students gaining enormous satisfaction from their research achievements and sharing with them the joys of publishing their data has been a particular pleasure. The use of active assessment strategies contextually adjusted to the research laboratory is a powerful way to learn what it is that students are learning and how the program can be refined to better achieve its research and education goals. Assessment of the type we propose is not construed as an extraneous intrusion into the education and research of student-researchers. As we have tried to show throughout this book, the type of assessment we developed that was situated within and attuned to a specific scientific inquiry process that we use in our laboratory directly enhances both our understanding of the educational benefits of our program and the quality of the student research that we produce. In this sense active assessment fulfills the aims of both good educational practice and quality control of undergraduate research.

At present, the research laboratory as an integrated environment for achieving educational goals is greatly underused. It is my and my collaborators hope that this book, with its explication of the basis for active assessment coupled with the PHIRE case study, will inspire numerous other scientist-educators to devise innovative programs that engage high school and undergraduate students in authentic scientific research. Our aim in writing this book was to make assessment relevant and accessible to a range of potential science educators. We believe that bringing students into the laboratory and assessing them closely in relation to specific scientific inquiry processes is the way forward. This is an approach that sees science as a specific, humanly valuable activity. We have experienced the excitement that students feel
when they are involved in real science and active assessment has allowed us to carefully assess what this learning actually means in conceptual, physical, and representational terms. We know that the students who graduate from our program are well versed in what conducting microbiological science means. Our hope is that this book will help scientists to become scientist-educators dedicated to the enhancement of science and science education through the processes of integrated science-based, education programs evaluated using active assessment procedures. Ultimately our belief is that real science inquiry processes and the development of valuable scientific knowledge must be at the center of educational and assessment practices in the sciences. If done properly this could result in a revolution in the way science education is conducted around the world. Just imagine 100,000 laboratory-based science education programs producing scientific knowledge utilizing even just 10 undergraduate student-researchers. This would change the face of how science is understood and developed around the world.
References


References


References


References


References


References


Index

Note: The locators with the letters ‘f’ and ‘t’ refers to figures and tables respectively

A
Active assessment
- book design, 8–9
- characteristics of, 4–6
- context/aims, 1–2
- definition, 3–4
- historical developments in science education, 2–3
- importance of active engagement, 6–8
- scientist-educator engagement and, 7
- student engagement in educational process, 7, 8
- scientists as educators, 1
- underpinning principles of active assessment, 4–6

Active assessment development
- process, schematic representation, 46f
- summary of main aspects, 52t–53t

Agarose gel electrophoresis of restriction-digested phage DNA., 65f

Aims of undergraduate phage hunting, 81–83

Analytical framework for defining scientific inquiry, 17–20
- development of assessment strategy for undergraduate scientific inquiry, 19
- scientific inquiry process (Hanauer), types of information, 17–18, 18f

ASIA, see Authentic scientific inquiry assessment (ASIA)

Assessing scientific inquiry
- active assessment development, 41–42
- approaches to, 31–37, 32f
- authentic assessment, 37–39
- characteristics of Authentic Scientific Inquiry Assessment (ASIA), 39–41

See also Scientific inquiry

Assessment
- aims of
- characteristics of educational program, 49
- content of scientific knowledge, 50
- questions, active assessment, 50
- developing assessment program, 26–29
- diagnostic, 24–25
- distinctions in elicitation methods, 25–26
- evaluation/testing, 23–24
- formative, 24–25
- of laboratory notebooks, rubric for, 107t
- instructor key, 108t–109t
- measures/knowledge types/timing/weighting of measures, 98t
- and science, 23
- summative, 24–25
- tools, process of piloting, 54

Authentic scientific inquiry assessment (ASIA), 45

B
Bacteriophage
- definition, 99–100
- evaluation of uniqueness of isolated, 96
- isolation/identification process, decision-making reasoning/calculation, see Decision-making reasoning/calculation, bacteriophage isolation/identification process
- selection of maximum yield, 96

Bacteriophage Institute of Pittsburgh, 7
130

BIO2010: Transforming Undergraduate Education for Future Research Biologists, 2

BLASTP, 70
search using BLAST server at NCBI, output of, 91f

D
Decision-making reasoning/calculation, bacteriophage isolation/identification process, 83–94
electrophoresis, 87–88
as a logical structure, 88
gene determination (annotation)
DNA Master, 89
Genemark program, 88
general principles of bacteriophage genomics, 90–94
Glimmer, 88–89
Glimmer program output, 90, 90f
graphic output of genemark program, 89, 89f
NCBI Gene BLASTing, 89
output of BLASTP search using BLAST server at NCBI, 91f
representations of genome analyses using DNA master, 92f
six-frame translation of a phage genome segment, 93f
phage amplification, measuring bacteriophage concentration for maximum yield, 86–87
visual analysis of web pattern on plate/logical structure, 86–87
phage isolation, identification through infection, 84–85
as logical structure, 85
schematic definition, 85
phage purification through selective infection, 85
schematic representation/logical structure, 85
Diversity of scientific inquiry
definition of scientific inquiry (Chinn and Malhotra), 12, 13–14
handbooks (manuals) in schools, analysis of, 12
levels and types of scientific inquiry, 14–15
See also Scientific inquiry
DNA
comparison, 96

Master, 68f, 89
restriction enzyme, 96
Double-stranded DNA (dsDNA), 56
Drug-resistant tuberculosis (XDR-TB), 72

E
Educational assessment strategy for scientific inquiry process of phage hunting, 94–97
knowledge presentation performance test, 97
notebook assessment tool, 96–97
physical checklist, 95
substantive knowledge test, 94–95
visual literacy test, 95–96
Electrophoresis, 87–88, 96
as logical structure, 88

G
Gene determination (annotation)
DNA Master, 89
Genemark program, 88
General principles of bacteriophage genomics, 90–94
Glimmer, 88–89
Glimmer program output, 90, 90f
graphic output of genemark program, 89, 89f
NCBI Gene BLASTing, 89
output of BLASTP search using BLAST server at NCBI, 91f
representations of genome analyses using DNA master, 92f
six-frame translation of phage genome segment, 93f
Gene identification, 96
Genemark program, 88
Genetic mosaicism, 71
Genome
analyses using DNA master, representations of, 92f
annotation and analysis, 96
map of mycobacteriophage L5, 70f
Glimmer, 88–89
program output, 90f
See also Gene determination (annotation)

H
Host range, 59
Index

I
Inquiry, scientific, 15–17
contextualized scientific knowledge, 15–16
development of personal and cultural knowledge, 15
progression toward high-order problem solving, 16
scientific inquiry as multi-stage and multi-representational process, 16–17
social interaction for scientific goals, 16
See also Scientific inquiry

K
K-12 culture, 78
Knowledge presentation performance test, 97
conference poster development, stages, 111–112
conference poster presentation, 112–113

L
The Laboratory Structure and Task Analysis Inventory, 12
Lysate concentration (titer), calculation of, 96

M
Microbiological technique of plating, checklist developed for, 101
Mycobacteriophage
barnyard, electron micrograph of, 57f
plaques on lawns of *M. smegmatis*, 58f

N
National Center for Biotechnology Information (NCBI), 70
National Science Education Standards, 2, 11, 24
NCBI, see National Center for Biotechnology Information (NCBI)
Notebook
assessment tool, 96–97
final assessment of laboratory notebooks, 107–110
knowledge presentation performance test, 110–113
on-going laboratory notebook usage and provision of feedback, 106–107
program beginnings and notebook introduction, 104–106
time periods, instructor’s role, 104
writing, principles of
principle of autonomous replication, 105
scientific argument, full presentation of, 105–106
Nucleotide sequence comparisons of phage genomes, 69f

P
PHIRE, see Phage Hunters Integrating Research and Education program (PHIRE)
Phage
amplification, 96
definition, 56
genome segment, six-frame translation of, 93f
hunting, educational assessment strategy for scientific inquiry process, 94–97
knowledge presentation performance test, 97
notebook assessment tool, 96–97
physical checklist, 95
substantive knowledge test, 94–95
visual literacy test, 95–96
hunting process, physical techniques, 95
identification through infection, 84–85
as logical structure, 85
schematic definition, 85
isolation, 96
measuring bacteriophage concentration to determine/produce maximum yield, 86–87
visual analysis of web pattern on plate/logical structure, 86–87
population, generating pure, 96
purification through selective infection, 85
schematic representation/logical structure, 85
Phage Hunters Integrating Research and Education program (PHIRE), 55
Phage Hunting Outreach Program, 81
Phamerator, 70
Phamily circles, 70, 71f
PHIRE assessment tools
knowledge presentation performance test
conference poster development, stages, 111–112
conference poster presentation, 112–113
notebook assessment tool
final assessment of laboratory notebooks (end of program), 107–110
knowledge presentation performance test, 110–113
on-going laboratory notebook usage and provision of feedback, 106–107
program beginnings and notebook introduction, 104–106
time periods, instructor’s role, 104
physical checklist, 100–101
substantive knowledge test, 99–100
visual literacy test, 101–103
templates, 102
test one: identification of bacteriophage, 103
wet-lab tasks, template, 102
visual literacy test one: identification of bacteriophage, 103

**PHIRE program (Phage Hunting Integrating Research and Education)**

**assessment strategy**

- aims of undergraduate phage hunting, 81–83
decision-making reasoning and calculation in bacteriophage isolation and identification process, 83–94
educational assessment strategy for scientific inquiry process of phage hunting, 94–97
knowledge presentation performance test, 97
notebook assessment tool, 96–97
physical checklist, 95
substantive knowledge test, 94–95
visual literacy test, 95–96
weighting and timing, 97–98

bacteriophages
dark matter of the biological universe, 59–60
ubiquitous and host specific, 58–59
failures, 77
goals of, 60
mentees and mentors, opportunities/responsibilities, 75–76
multiple successes, 76–77
name of the program, 73–74
parallel projects, pros and cons, 74–75
phages, definition, 56–58
plaque assay, 57
power of discovery, 60–61
qualifications, 74
scheduling and flexibility, 76

seven attributes of
attribute 1: technical simplicity with transition to complexity, 78
attribute 2: conceptual simplicity with transition to complexity, 79
attribute 3: flexible scheduling compatibility, 79
attribute 4: multiple achievement milestones, 79
attribute 5: parallel projects and mentoring, 79
attribute 6: authentic publishable research, 79
attribute 7: project ownership, 79
ten-step program, 61f
agarose gel electrophoresis of restriction-digested phage DNA, 65f
DNA Master program, 67
equilibrium density gradient purification of mycobacteriophage particles, 63f
nucleotide sequence comparisons of phage genomes, 69f
step 1: phage isolation, 61–62
step 2: phage purification, 62
step 3: phage amplification, 62–63
step 4: electron microscopy, 63–64
step 5: nucleic acid extraction and restriction analysis, 64–66
step 6: DNA sequencing, 66–67
step 7: genome annotation, 67
step 8: comparison of DNA sequence to known phage genomes, 67–69
step 9: comparative genomic analysis, 69–71
step 10: publication, 72
ten steps of, 61f
transitions: from concrete beginnings to abstract representation, 77–78

Physical checklist, 95, 100–101
Piloting, assessment, 54
Plaque
- assay, 57–58
- identification of, 96
Plaque-forming unit, isolation of, 58, 60

R
RED, see Restriction enzyme digestion (RED)
Restriction digestion, 64
Restriction enzyme digestion (RED), 87
Rubric development, 51–54
Science inquiry
- active assessment development, stages of assessment piloting, 54
- schematic representation of process, 46–49
- stage 1: empirical description of scientific inquiry, 46–49
- stage 2: definition of aims, 49–50
- stage 3: tool development, 50–51
- analytical framework for defining, 17–20
- development of assessment strategy for undergraduate scientific inquiry, 19
- scientific inquiry process (Hanauer), types of information, 17–18, 18f
- assessing
  - empirical description of, 47f
  - questions, according to two axes, 47–49

Scientific teaching, 2
- Scientist-educator, reflections from integrating missions, 116–117
- integrating the missions, 116–117
- practical concerns
  - assessment, 118–119
  - money and time, 118
  - program attributes revisited, 119–120
  - space, 117–118

Single-stranded DNA (ssDNA), 56
- Substantive knowledge test, 94–95, 99–100

Scientific inquiry capabilities (Zachos), 33

Scientific teaching, 2
- Scientist-educator, reflections from integrating missions, 116–117
- integrating the missions, 116–117
- practical concerns
  - assessment, 118–119
  - money and time, 118
  - program attributes revisited, 119–120
  - space, 117–118

Single-stranded DNA (ssDNA), 56
- Substantive knowledge test, 94–95, 99–100

Visual literacy test, 95–96, 101–103
- identification of bacteriophage, 103f
- templates, 102
- test one: identification of bacteriophage, 103
- wet-lab tasks, template, 102

XDR-TB (drug-resistant tuberculosis), 76
Color Plates

Fig. 2.2 A Schematic Representation of the Scientific Inquiry Process

Fig. 5.2 A schematic representation of the scientific inquiry process
Color Plates

Fig. 6.3 The ten steps of the PHIRE program. Each of the ten steps in the PHIRE program are illustrated with the central arrow showing the transition from concrete to abstract comprehension from step 1 to step 10.

Fig. 6.6 DNAMaster: A program for phage genome annotation. Screen shots are shown from several representative windows of DNAMaster (A–E). (A) Window presenting features of mycobacteriophage BPs. The “Feature” tab is selected, showing a list of the genes with their coordinates. (B) The “Compare Genomes” window showing the results of a BlastP search of phage BPs gp7. (C) Comparison of phage BPs and phage Halo genomes, with each yellow box indicating the position of genes that are shared between the two genomes. (D) The “Frames” window showing the positions of stop and start codons in all six reading frames. The positions of identified opening reading frames are colored either green (forward direction) or red (reverse direction). (E) Plot of base composition showing the average GC% content across the BPs genome. These represent just a small subset of the available functions with DNAMaster (available as a free download from http://cobamide2.bio.pitt.edu)
Color Plates

Fig. 6.8 Genome map of mycobacteriophage L5. The L5 genome is represented as a horizontal bar with markers spaced at 1 kbp intervals. Each of the predicted genes is shown as a colored box either above or below the genome; those above are transcribed rightward, and those below are transcribed leftward. The map was generated by the program Phamerator, which sorts the predicted genes into phamilies (Phams) as functions of their predicted amino acid sequence relatedness. The Pham number is shown above each gene with the number of Pham members in parentheses, and each gene is colored according to its Pham. Some of the predict gene functions are noted.

Fig. 6.9 Phamily circles. Phamily circles for Pham38 and Pham41 are shown. In the L5 genome (see Fig. 6.8), genes 58 and 59 are members of these two Phams, respectively, and the phamily circles suggest that they have distinct and different evolutionary histories. Each of the phages is listed around the circumference of each circle and an arc is drawn between members of each phamily with line thickness corresponding to strength of the relationship.
Fig. 7.1 Graphic Output of GeneMark Program. A segment of phage Adjutor genome sequence was analyzed using the program GeneMark, which generates a graph of the coding potential in each of the six possible open reading frames (ORFs) (labeled 1–6; frames 1–3 are expressed in the rightward direction, and 4–6 in the reverse direction). ORFs are identified as horizontal lines between translation stop codons (down ticks), and potential start codons are shown as up ticks. Coding potential is indicated by the line graph. The thick horizontal gray lines (still italicized) are considered areas of interest. One such gene is highlighted in green.
**Color Plates**

Fig. 7.2 Glimmer Program Output. A segment of the phage Adjutor genome sequence was analyzed by Glimmer generating a list of potential ORFs, their start and stop codons, reading frame, and evaluative score. The red box designates one specific candidate gene.
Color Plates

Fig. 7.3 Output of a BLASTP search using the BLAST server at NCBI. The amino acid sequence of a predicted phage protein was submitted as a search query to the server, which returned a summary of the search results; a screen shot of the results window is shown. At the top, a diagram shows the correspondence of a putative conserved domain identified in the search with a linear representation of the query sequence. Below that is a diagram of related sequences identified in the search, with the closest matches shown in red. The identities of the specific matches and their scores are shown in the bottom panel.
Fig. 7.4 Representations of genome analyses using DNA Master. (A) The output from the “Frames” routine in DNA Master showing six reading frames, with start *(short vertical lines)* and stop codons *(full height lines)* indicated. (B) ORFs with strong prediction for coding potential predicted by GenMark and Glimmer are designated by a *thick green horizontal line*. (C) Clicking on a segment of a frame of interest produces a *thin green* line and also links to the “choose start” routine shown in (D) the ORF and start codon corresponding to the frame segment in (C) is in *red*. (4). (E) The genome segment highlighted in (C) is revealed as a selected segment *(black background)* in the “Sequence” panel of DNA Master, and can be added to the annotation or copied as either a nucleotide or a translated amino acid sequence for BLAST analyses.
Fig. 7.5 Six-frame translation of a phage genome segment. DNAMaster provides an output showing the DNA sequence of both strands as well the amino acid sequences in each of the potential six reading frames. The amino acid sequences of the potential ORFs predicted by Glimmer and GenMark are colored. Each of the potential start codons shown in Fig. 7.4E (and within this DNA segment) are shown boxed.