2015

An Examination of Contingency in Synthetic Genomics Research and Implications for National Security

Fye-Marnien, Shannon

http://hdl.handle.net/1920/10188

a service of Mason Publishing
AN EXAMINATION OF CONTINGENCY IN SYNTHETIC GENOMICS RESEARCH AND IMPLICATIONS FOR NATIONAL SECURITY

by

Shannon R. Fye-Marnien
A Dissertation
Submitted to the
Graduate Faculty of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree of
Doctor of Philosophy
Biodefense

Committee:

______________________________ Sonia Ben Ouagrham-Gormley, Chair
______________________________ Serguei Popov
______________________________ A. Trevor Thrall
______________________________ Kathleen M. Vogel
______________________________ Gregory Koblentz, Program Director
______________________________ Mark J. Rozell, Dean

Date: _________________________ Fall Semester 2015
George Mason University
Fairfax, VA
An Examination of Contingency in Synthetic Genomics Research and Implications for National Security

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Shannon R. Fye-Marnien
Master of Science
George Mason University, 2006
Bachelor of Science
California State University, Sacramento, 2004

Director: Sonia Ben Ouagrham-Gormley, Chair, Professor
Biodefense Program

Fall Semester 2015
George Mason University
Fairfax, VA
DEDICATION

This dissertation is dedicated to my beloved grandmother, Roberta Corbett Huffman, who is greatly missed.
ACKNOWLEDGEMENTS

This dissertation is the result of the encouragement and insight of many people. First and foremost, I would like to thank my committee members, Sonia Ben Ouagrham-Gormley, Kathleen Vogel, Serguei Popov, and Trevor Thrall, for providing input that has improved this project immensely. I am deeply honored to have worked with each of them. Thanks especially to my committee chair, Sonia Ben Ouagrham-Gormley, for her guidance and for helping me publish some of my research findings in a peer-reviewed journal (Sonia Ben Ouagrham-Gormley and Shannon R. Fye, “Restricted Science,” Frontiers in Public Health 2, no. 158 [2014]).

This work would not have been possible without the researchers who allowed me to interview them. Many thanks to all of the people I interviewed, including Dr. John Glass, Dr. Ray Chuang, Dr. Tim Gardner, Dr. Brian Nosek, and several other researchers who wished to remain anonymous: All provided me with extremely valuable information.

Thanks to the coworkers who have encouraged me, especially Anton Jareb, Dr. Denise Brown-Anderson, Dr. Steve Charbonneau, Carie Mullins, Elaine Gresham, Phil Smith, Raphael Perrino, Jason Hay, Carissa Christensen, Kathy Schvaneveldt, Dr. Shawn “Froggi” Jackson, and Dr. Jerome Holton.

I am also grateful for all of the excellent professors I’ve had over the years whose passion for teaching left a lasting impression on me, especially Dr. Susanne W. Lindgren at California State University, Sacramento. Finally, I’m grateful to my parents, grandparents, husband, and friends for their love and encouragement.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xi</td>
</tr>
<tr>
<td>Abstract</td>
<td>xii</td>
</tr>
<tr>
<td>Chapter 1 Identifying the Potential Threats of Synthetic Genomics</td>
<td>1</td>
</tr>
<tr>
<td>Overview of Findings</td>
<td>5</td>
</tr>
<tr>
<td>Methodology</td>
<td>8</td>
</tr>
<tr>
<td>Dissertation Organization</td>
<td>11</td>
</tr>
<tr>
<td>Chapter 2 Synthetic Biology and Synthetic Genomics: Understanding the</td>
<td>13</td>
</tr>
<tr>
<td>Technology, Current Concerns, and Threat-Prevention Policies</td>
<td></td>
</tr>
<tr>
<td>What are Synthetic Genomics and Synthetic Biology?</td>
<td>14</td>
</tr>
<tr>
<td>Discourse on Biotechnological Advances and National Security</td>
<td>27</td>
</tr>
<tr>
<td>Technological Determinism</td>
<td>35</td>
</tr>
<tr>
<td>Inevitability of Nefarious Use</td>
<td>38</td>
</tr>
<tr>
<td>Theoretical Possibilities</td>
<td>43</td>
</tr>
<tr>
<td>Threat-Mitigation Policies: Restricting Access to Dangerous Pathogens</td>
<td>46</td>
</tr>
<tr>
<td>and Associated Genetic Material</td>
<td></td>
</tr>
<tr>
<td>The Select Agent Rule</td>
<td>47</td>
</tr>
<tr>
<td>The Australia Group</td>
<td>49</td>
</tr>
<tr>
<td>Export Administration Regulations</td>
<td>51</td>
</tr>
<tr>
<td>Biological and Toxin Weapons Convention</td>
<td>52</td>
</tr>
<tr>
<td>Chemical Weapons Convention</td>
<td>53</td>
</tr>
<tr>
<td>Proposed New Framework to Address the Potential Threats of DNA Synthesis</td>
<td>56</td>
</tr>
<tr>
<td>Screening Genetic Sequences</td>
<td>58</td>
</tr>
<tr>
<td>Conducting Institutional Reviews of Synthetic Biology Experiments</td>
<td>64</td>
</tr>
<tr>
<td>Censoring Scientific Literature</td>
<td>70</td>
</tr>
</tbody>
</table>
Appendix D: IGSC Harmonized Screening Protocol ....................................... 248
Appendix E: Interview Questions for Gene Synthesis Firms ............................ 254
References ............................................................................................................ 256
LIST OF TABLES

Table                                      Page
Table 1. Summary of gene synthesis firm interviews......................................................... 10
Table 2. Summary of JCVI interviews. ................................................................. 11
Table 3. Summary of U.S. regulatory framework and shortfalls............................................. 55
Table 4. Gene synthesis company methods to resolve contingencies. ............................ 144
Table 5. Summary of contingencies and resolution methods. ........................................... 174
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. Oligonucleotide, gene, and DNA synthesis.</td>
<td>16</td>
</tr>
<tr>
<td>Figure 2. DNA synthesis cost over time.</td>
<td>18</td>
</tr>
<tr>
<td>Figure 3. Productivity in DNA synthesis over time.</td>
<td>20</td>
</tr>
<tr>
<td>Figure 4. A word cloud showing common words used in threat assessments on synthetic genomics and synthetic biology.</td>
<td>42</td>
</tr>
<tr>
<td>Figure 5. Steps in the gene assembly process.</td>
<td>108</td>
</tr>
<tr>
<td>Figure 6. Example of an ideal-looking DNA sequence.</td>
<td>128</td>
</tr>
<tr>
<td>Figure 7. Example of a chromatograph with a run on nucleotides.</td>
<td>130</td>
</tr>
<tr>
<td>Figure 8. Contingencies in the gene synthesis process.</td>
<td>131</td>
</tr>
<tr>
<td>Figure 9. Graphic representation of the JCVI’s synthesis of the <em>M. mycoides</em> genome.</td>
<td>157</td>
</tr>
<tr>
<td>Figure 10. User question about Gibson Assembly™ on Reddit’s Labrats page.</td>
<td>186</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

Base pair........................................................................................................... bp
Kilobase ............................................................................................................. kb
Megabase .......................................................................................................... Mb
Nucleotide ........................................................................................................ nt
ABSTRACT

AN EXAMINATION OF CONTINGENCY IN SYNTHETIC GENOMICS RESEARCH AND IMPLICATIONS FOR NATIONAL SECURITY

Shannon R. Fye-Marnien, Ph.D.

George Mason University, 2015

Dissertation Director: Dr. Sonia Ben Ouagrham-Gormley, Chair

The fields of synthetic genomics and synthetic biology have garnered much attention in the biodefense community and the general public due to their dual-use nature: their potential use for both peaceful applications and harmful purposes. Although many researchers are using the technologies to improve medical diagnostics, prophylaxes, and therapeutics, several publications have prompted concern over the technologies’ national security implications. These publications include a 2005 article that describes how researchers recreated the influenza virus that caused a pandemic in 1918, and a 2006 article in The Guardian that describes how a reporter easily ordered part of the smallpox genome from a synthetic genomics firm. These publications have spurred fears that terrorists or other malefactors could use synthetic genomics and synthetic biology to create dangerous pathogens. However, many mainstream assessments that voice these concerns assume that the results achieved in one experiment can be easily replicated if
the necessary materials and protocols are provided, and that advances in synthetic genomics and synthetic biology will reduce the level of skill required to use the technologies. This runs counter to research in the field of science and technology studies, which indicate that there is a great deal of experimental contingency—unexpected technical difficulties—associated with using biotechnologies. Much of this contingency is due to the inherent limitations of working with biological systems, which are unpredictable and sensitive to their environments, as well as problems in reproducing well-established laboratory tasks in new contexts or settings. However, no study to date has evaluated the technical difficulties associated with synthetic genomics, a critical enabling technology for synthetic biology. This dissertation aims to fill this gap. I have interviewed representatives of gene synthesis firms and conducted a case study of the J. Craig Venter Institute’s synthesis of the *Mycoplasma mycoides* genome. These analyses will (1) illuminate what difficulties are associated with the use of synthetic genomics, and how those difficulties can affect the users of the synthesized DNA during their larger synthetic biology experiments, and (2) determine under what conditions malefactors could overcome those technical difficulties. In addition to enriching the body of literature that describes the level and type of expertise required to perform tasks in the life sciences, this dissertation ultimately aims to provide security analysts and government officials with better tools to improve threat assessments on biotechnologies, which might aid the development of more effective measures to counter perceived threats from dual-use technologies.
CHAPTER 1
IDENTIFYING THE POTENTIAL THREATS OF SYNTHETIC GENOMICS

The fields of synthetic genomics and synthetic biology have garnered much attention in the biodefense community and the general public due to their dual-use nature: their potential use for both peaceful applications and harmful purposes.1 Synthetic genomics is a field of biotechnology that produces genes or whole genomes chemically for use in a variety of scientific experiments, including synthetic biology.2 Synthetic biology is a field of biotechnology that involves the design and construction of novel biological pathways, organisms, or devices—or the redesign of existing biological systems—using genes or whole genomes that have been produced chemically via synthetic genomics.3 Although many researchers are using the technologies to improve medical diagnostics, prophylaxes, and therapeutics, several publications have prompted concern over the technologies’ national security implications.4 These include a 2002 study published in Science that describes the synthesis of the poliovirus from genes that

---

3 There is no standard definition for synthetic biology. The definition here is an amalgamation of the one provided in the following source: “What Is Synthetic Biology?,” Synbioproject.org, accessed January 24, 2015, http://www.synbioproject.org/topics/synbio101/definition.
were chemically synthesized; another *Science* article in 2005 that describes how researchers recreated the influenza virus that caused a pandemic in 1918; and a 2006 article in the United Kingdom’s *The Guardian*, in which a reporter ordered part of the smallpox genome from an unnamed synthetic genomics firm for approximately $40, ostensibly proving that it would be easy for terrorists to order synthetic genes to create a biological agent.

These and other studies spurred fears that terrorists or other malefactors could use synthetic genomics and synthetic biology to create pathogens from scratch. Terrorists could accomplish this in one of two ways: (1) by synthesizing a pathogen’s entire genome using commercially available, automated synthesizers or (2) by ordering individual pathogen genes from gene synthesis firms and then assembling these genes into a complete genome. In response to these concerns, national security analysts have proposed a number of policy measures aimed at decreasing the likelihood that synthetic genomics and synthetic biology are misused. These measures include requiring firms that synthesize genes to screen each order to ensure the sequences are not analogous to sequences found in controlled pathogens; requiring institutional review boards to review

---

8 Title 42 Code of Federal Regulations, Part 73 defines biological agents as any microorganism (including, but not limited to, bacteria, viruses, fungi, rickettsiae, or protozoa), or infectious substance, or any naturally occurring, bioengineered, or synthesized component of any such microorganism or infectious substance, that is capable of causing death, disease, or other biological malfunction in a human, an animal, a plant, or another living organism; deterioration of food, water, equipment, supplies, or material of any kind; or deleterious alteration of the environment.
synthetic genomics and synthetic biology experiments to ensure they do not present biosafety threats,⁹ and censoring scientific literature so that the methods sections for experiments of concern are obscured or omitted entirely.¹⁰

However, many of the current assessments that voice these security concerns assume that the results achieved in one experiment can easily be replicated if the necessary materials and protocols are provided. They also assume that advances in synthetic genomics and synthetic biology will reduce the level of skill required to use these technologies. This runs counter to research in the field of science and technology studies (STS), which shows that there is a great deal of experimental contingency—unexpected technical difficulty—that arises during life sciences research.¹¹ Several scientists have echoed these findings by describing the technical challenges synthetic biologists and other practitioners face.¹² Much of this contingency is due to (1) the unpredictable nature of deoxyribonucleic acid (DNA) and living organisms, which mutate and are sensitive to their environments; and (2) problems with reproducing well-established laboratory tasks in new contexts or settings. Although some of these

---

⁹ Garfinkel et al., “Synthetic Genomics.”
contingencies stem from ambiguous language in laboratory protocols and other forms of explicit knowledge,\(^{13}\) much of the expertise required in laboratory experiments is based on tacit knowledge\(^ {14}\) and skills that cannot be easily translated into words because they are based on motor skills and other hands-on practices that escape codification.

While there is rich data in the STS literature that describes the contingencies synthetic biologists encounter, there has been limited data that has examined the technical difficulties associated with synthetic genomics, a critical enabling technology for synthetic biology. How easy would it be for terrorists or other malefactors to synthesize a pathogen’s genome using published protocols and commercially available technologies? Could they transfer these synthetic genomes into cells using synthetic biology? What kind of skills would they need? What other factors might enable or prevent them from achieving successful results? To answer these questions, I interviewed four representatives from three U.S. gene synthesis firms to determine the extent to which they encounter technical difficulties. These firms specialize in producing gene-length DNA and are comprised of experts in the process of gene synthesis. I also conducted a case study of work conducted by the J. Craig Venter Institute (JCVI) to synthesize the large, 1.08 million base pair *Mycoplasma mycoides* genome, which was assembled in a step-wise process by combining small DNA fragments provided by gene synthesis firms into ever-larger fragments until the entire genome was assembled. In addition to identifying the difficulties associated with performing synthetic genomics, my interviews

\(^{13}\) Explicit knowledge is knowledge that is codified in the form of scientific protocols, instructions, software, and procedures.

\(^{14}\) Tacit knowledge is unarticulated knowledge acquired through hands-on experience, which involves either learning by directly performing a task or by observing and collaborating with colleagues.
also elucidate how synthesis firms and JCVI researchers overcome contingencies. This information helps to identify the challenges that malefactors might face in attempting such experiments and how they might resolve any difficulties they encounter.

The purpose of my research is not to prove or disprove that synthetic genomics is a threat to national security. Rather, the purpose of my research is to add to the body of literature in the STS field that describes the level and the type of expertise required to perform tasks in the life sciences and to offer a more nuanced evaluation of the technical skills that are required to exploit synthetic genomics and synthetic biology—a major failing of mainstream threat assessments. Since the potential security implications associated with synthetic genomics and synthetic biology have been widely discussed in the biodefense community, the outcome of this research can be used to inform discussions about the risks associated with emerging biotechnologies.

**Overview of Findings**

My interviews with gene synthesis firms show that terrorists or other malefactors would encounter problems at key stages in the gene synthesis process: when designing genes prior to synthesis, synthesizing genes that are known to be inherently problematic, assembling larger genes, cloning genes into vectors, and verifying that the genes they have synthesized contain the correct sequence. To overcome contingencies, synthesis firms primarily rely on standard operating procedures (SOPs). However, the SOPs are derived from in-house knowledge that is not publicly available. Furthermore, SOPs alone do not provide all of the information necessary to enable an individual to synthesize genes. New employees, for example, must observe the routines of more experienced scientists to learn how to synthesize genes, a process that can take as long as six months.
Although much of the work in commercial synthesis is automated and based on routinized procedures, the field is not completely automated and some of the work still requires human judgment and skill. Furthermore, automated systems and software cannot always prevent or resolve difficulties in the gene synthesis process, and they can introduce new problems that require human expertise to resolve. Indeed, nearly every aspect of gene synthesis depends on human knowledge—much of which is tacit—including the knowledge of which tasks should be performed by automated systems.

Previous research has shown that gene synthesis can be too difficult for many scientists to carry out, so they often outsource synthesis to commercial firms. Considering the challenges faced by these firms, who are comprised of experts in the field, terrorists with considerably less expertise are likely to encounter significant problems creating pathogen genes and genomes.

The interviews I conducted with JCVI researchers involved in the *M. mycoides* project show that terrorists or other malefactors would likely encounter problems assembling and purifying large genes, transplanting synthetic genomes into recipient cells, using polymerase chain reaction (PCR), and verifying that gene sequences are correct. Many of these difficulties are due to the unpredictable nature of DNA, sensitivities of host cells, and problems handling large fragments of DNA. All of the contingencies the JCVI researchers encountered required tedious, time-consuming trial and error to overcome. Although they used the academic literature and other published protocols as a starting point for resolving many of the difficulties they encountered, they

---

had to adapt the protocols to suit their specific needs—a time-consuming process. In addition, some aspects of the *M. mycoides* project required significant motor skills that could only be acquired through frequent repetition, such as pipetting the fragile *M. mycoides* genome gently enough to prevent it from shearing. This shows that the project was highly dependent upon JCVI researchers’ specialized—and sometimes tacit—expert knowledge.

The JCVI team included dozens of highly motivated experts who operated with all the necessary financial and material resources, yet they still encountered numerous difficulties, demonstrating that genome synthesis is hardly a straightforward or inexpensive process. Terrorists and other malefactors who are trying to assemble a pathogen genome would likely face even more challenges, especially if they lack the appropriate expertise and resources and if the genome is large and complex. Currently, the experiment is specific to *M. mycoides*, and the JCVI researchers who developed the method have not been able to adapt it for use with other organisms. Therefore, terrorists would not be able to use the JCVI researchers’ protocol to create a select agent.

Thus, my interviews indicate that synthesizing genomes is a complex undertaking that is far from being automated. As such, terrorists are unlikely to master genome synthesis. Further, published protocols are often not detailed enough for even experienced scientists to use. Therefore, they are not simple recipes that terrorists can use to successfully replicate a previous experiment. Because the process of synthesizing genomes is complex and published protocols are insufficient, researchers rely on specialized and tacit knowledge to overcome experimental contingencies. Given that this
knowledge is difficult to acquire and transfer, most mainstream threat assessments have overstated the threats arising from synthetic genomics because they have not adequately considered the skills required to use the technology. Instead, these assessments emphasize terrorists’ intent to develop biological agents and assume that their capability to do so merely involves acquiring materials and protocols. These findings point to a need for more rigorous threat assessment methodologies, such as those that incorporate STS input to characterize the sociological and organizational factors that influence the way technologies are developed and used. They also point to the need for federal advisory bodies and international nonproliferation policies to consider the role of tacit knowledge in developing and using biotechnologies.

Although my research demonstrates that terrorists or other malefactors would encounter significant technical challenges in using synthetic genomics and synthetic biology to create a pathogen, a persistent malefactor could eventually acquire the tacit knowledge necessary to achieve this goal. Therefore, the synthetic genomics and synthetic biology communities should embrace and improve measures meant to prevent this misuse, such as current methods to screen sequences that are ordered from gene synthesis firms. The U.S. government should also improve institutional review of dual-use research.

Methodology
To answer my research questions, I began by examining the current discourse on biotechnological advances like synthetic genomics and synthetic biology to identify key assumptions made in threat assessments of the technologies. I then examined the current U.S. regulatory framework to determine if it adequately addresses the perceived threats.
also reviewed the STS literature to identify contingencies associated with life sciences research and the means by which individuals and organizations acquire the tacit knowledge and expertise that help them resolve contingencies. The knowledge I gained from my review of the STS literature formed the analytical framework for my interviews with the scientists who conduct synthetic genomics and synthetic biology.

To identify the technical difficulties associated with gene synthesis and the means by which these difficulties are overcome, I conducted interviews with representatives from gene synthesis firms in May, June, and July of 2011. I also conducted a follow-up interview in August 2013 with one of the companies to ensure the original interview information was still valid. I chose to obtain data through interviews because data about industrial gene synthesis is proprietary and therefore not found in current publications. Due to the gene synthesis industry’s highly competitive environment and the relative scarcity of firms based on the United States, only three companies were willing to participate in my research. All three companies are headquartered in the United States, and all are similar in size. One firm specializes in synthesizing short fragments of up to 50 kilobases (kb), while the other two specialize in larger sequences of up to 250 kb.

All interviewees were managers at their respective companies. However, all had scientific backgrounds in molecular biology and genomics and were familiar with the process of gene synthesis. Due to proprietary concerns, all interviewees asked to remain anonymous. In this dissertation, these anonymous interviewees will be cited as follows: Interview with employee 1 of gene synthesis company A, date. Subsequent references will be cited as follows: Company A, interview; dates will only be specified in the
subsequent references if there were multiple interviews. Table 1 summarizes the number of people I interviewed from each company, as well as the total number of interviews I conducted with each company.

Table 1. Summary of gene synthesis firm interviews.

<table>
<thead>
<tr>
<th>Company designation</th>
<th>Number of people interviewed</th>
<th>Total number of interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company A</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Company B</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Company C</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

To identify the technical difficulties associated with large-gene and genome assembly, I interviewed two JCVI researchers who participated in the *M. mycoides* project and another who participated in the assembly of the *M. genitalium* genome. All interviewees were doctoral-level researchers with at least 15 years of laboratory experience in molecular biology, biochemistry, and related fields. I conducted an in-person interview at JCVI’s Rockville, Maryland, campus on March 4, 2013 with Dr. John Glass, leader of JCVI’s Synthetic Biology and Bioenergy Group. I conducted two telephone interviews with Dr. Ray-Yuan Chuang, Assistant Professor in JCVI’s Synthetic Biology Group. The first interview was on September 19, 2013 and the second interview was on February 12, 2015. This second interview was conducted primarily to determine whether the researcher had overcome a technical difficulty that was described in the previous interview. Finally, I conducted one telephone interview on September 26, 2013 with a JCVI molecular biologist who requested anonymity. Dr. Chuang and the anonymous researcher participated in aspects of genome assembly, while Dr. Glass
oversaw genome transplantation. More recently, Dr. Chuang also participated in transplantation. Table 2 summarizes the number of interviews I conducted, and what aspects of the experiment each interviewee was involved in.

<table>
<thead>
<tr>
<th>Researcher interviewed</th>
<th>Number of interviews</th>
<th>Interview location</th>
<th>Role in <em>M. mycoides</em> experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. John Glass</td>
<td>1</td>
<td>Rockville, MD</td>
<td>Assembly</td>
</tr>
<tr>
<td>Dr. Ray Chuang</td>
<td>2</td>
<td>Telephone</td>
<td>X</td>
</tr>
<tr>
<td>Anonymous researcher</td>
<td>1</td>
<td>Telephone</td>
<td>X</td>
</tr>
</tbody>
</table>

I complemented the data collected from interviews by reviewing presentations and papers written by scientists I was unable to interview. I also searched online forums to identify additional evidence of contingencies in using specific biotechnologies. I then compared my findings with those from previous STS case studies on biotechnologies, contrasted them with the mainstream security threat assessment discourse, and described their national security implications.

**Dissertation Organization**

The dissertation proceeds as follows. In chapter 2, I discuss the dual-use nature of synthetic genomics, analyze what the current threat assessments say about biotechnologies, such as synthetic genomics, and provide an overview of current policies aimed at preventing its misuse. Chapter 3 summarizes the STS literature, which identifies the experimental contingencies associated with biotechnology research that most threat assessments fail to characterize. The chapter also describes how malefactors could acquire the knowledge that might allow them to resolve any experimental contingencies.
that they encounter. Using information from my interviews with gene synthesis firms, I characterize malefactors’ ability to synthesize genes in chapter 4. In chapter 5, my case study of the JCVI’s M. mycoides project characterizes malefactors’ ability to assemble large genes or genomes from genes provided by synthesis firms. I summarize the major findings from my interviews and compare them to previous STS research in chapter 6. Lastly, I conclude my research in chapter 7 by offering overarching recommendations for conducting more nuanced threat assessments, incorporating a sociotechnical frame into national biosecurity initiatives and international nonproliferation policies, improving measures to prevent the misuse of synthetic genomics and synthetic biology, and engaging with other researchers to obtain data on experiment reproducibility.
CHAPTER 2
SYNTHETIC BIOLOGY AND SYNTHETIC GENOMICS: UNDERSTANDING THE TECHNOLOGY, CURRENT CONCERNS, AND THREAT-PREVENTION POLICIES

Dual-use technologies have the potential to be used for both peaceful and harmful purposes. Many national security analysts are concerned that terrorists or other malefactors could use synthetic biology and its enabling technologies, including synthetic genomics, for hostile purposes or that scientists performing synthetic biology research would accidentally release synthetic organisms into the environment, negatively affecting human health or the environment.¹ This dissertation examines the former concern: that terrorists or other malefactors could use synthetic biology and its enabling technologies for hostile purposes.

Security analysts are concerned about biotechnologies like synthetic genomics and synthetic biology because they are undergoing rapid advancement and associated materials and techniques are becoming easily accessible worldwide. In addition, security analysts often cite decreasing costs and increasing automation as evidence that

malefactors will be able to carry out attacks using biological agents obtained via synthetic genomics. However, these security assessments fail to factor in the human skill required to use biotechnologies, assuming that access to materials and procedures is sufficient to achieve successful results. STS literature shows that scientists of all skill levels—including experts—encounter a variety of difficulties that require extensive troubleshooting when using biotechnologies, even those that have been used widely for decades and are believed to be almost entirely automated. The vast majority of threat assessments omit this information, resulting in ineffective or unnecessary regulations in nonproliferation and technology policies.

In this chapter, I start by providing an overview of synthetic genomics and synthetic biology. I then analyze the current discourse on biotechnological advances, synthetic biology, and synthetic genomics and examine several key assumptions made in current threat assessments. Next, I review the regulatory framework that aims to prevent the spread of sensitive biological technologies, and I highlight some of the newly suggested rules aimed at addressing the specific threats emanating from synthetic biology and synthetic genomics. I conclude by highlighting the strengths and weaknesses of these regulations in mitigating the risk that synthetic genomics could be used for nefarious purposes.

**What are Synthetic Genomics and Synthetic Biology?**

Synthetic genomics is the production of gene- or genome-length deoxyribonucleic acid (DNA) from chemically derived oligonucleotides, which are short fragments of
DNA that are often single-stranded and typically no longer than 50 nucleotides long. Prior to the emergence of the first commercial DNA synthesizers in the early 1980s, scientists extracted DNA of interest using conventional genetic engineering techniques, which were time-consuming. But with advances in biotechnology, researchers can circumvent the tedious standard genetic engineering methods of obtaining DNA from a host organism and instead obtain synthetic DNA relatively easily from a machine. While these machines are easier to use than traditional genetic engineering methods, it can still be complex to synthesize DNA, so researchers typically outsource the synthesis to commercial firms. These firms can deliver the requested DNA to customers within days, depending on the length of the DNA sequence ordered.

There are three terms that describe the services these companies provide: oligonucleotide synthesis, gene synthesis, and DNA synthesis. Oligonucleotide synthesis is the chemical production of small fragments of DNA that are single-stranded and not functional. These smaller fragments are typically combined into larger fragments in laboratories by their end users. Gene synthesis—sometimes also called “synthetic genomics”—is the assembly of larger gene- or genome-length DNA fragments (typically more than 200 nucleotides in length) that are functional and double-stranded, using chemically produced oligonucleotides. DNA synthesis is an umbrella term that describes the method by which nucleic acids—the building blocks of DNA—are chemically

---

2 A nucleotide (nt) is one of the structural components of DNA. A nucleotide consists of a base (either adenine, thymine, guanine, or cytosine), a molecule of sugar, and a molecule of phosphoric acid.
3 Marsic et al., “PCR-Based Gene Synthesis” (see chap.1, n. 14).
synthesized in vitro without initial template DNA. Since both oligonucleotide and gene synthesis require DNA, both are types of DNA synthesis. Figure 1 summarizes these three terms and their relationship to each other. This dissertation focuses on gene synthesis because, unlike oligonucleotide synthesis, gene synthesis is not entirely automated and because oligonucleotide synthesis does not directly produce functional genes that can be combined into an entire genome.

![Figure 1. Oligonucleotide, gene, and DNA synthesis.](image)

The convenience and low cost of DNA synthesis has driven substantial growth and competition, forcing prices so low that profit margins are minimal. One
oligonucleotide synthesis company, which claims to have the lowest prices in the industry, advertises that it can synthesize DNA for as little as $0.23 per base pair (bp).⁵

Competition in the gene synthesis industry, especially, is fierce because assembling larger DNA fragments is still technologically challenging, and companies are working to develop methods that are faster and more accurate than those of their competitors. There are fewer than ten DNA synthesis companies in the United States and only dozens worldwide.⁶ These companies use technologies and methods that are proprietary. Research by industry expert Rob Carlson of Bioeconomy Capital Research, shown in figure 2, reveals that synthetic gene prices reached their lowest point in 2012 and have been stable since, with no further drops anticipated in the near future unless there is another round of technical innovation or a surge in demand.⁷

Due to the minimal profit margins, synthesis companies attempt to differentiate themselves by the length of genes they can synthesize, the ease of online ordering, and the speed of order turnaround. These companies also offer services meant to improve and validate the expression of the synthesized genes in target vectors, such as gene or codon⁸ optimization. Ordering synthetic products is simple: for most orders, customers specify their target sequence and vector online, and orders are delivered within days.

⁶ Carlson, Biology is Technology (Harvard University Press, 2010).
⁸ A codon is a sequence of three nucleotides that together code for a specific amino acid. For example, the codon UAU (where U is the nucleotide uracil and A is the nucleotide adenine) codes for the amino acid tyrosine.
Much of the speed associated with synthesis is due to the automation of associated equipment. Parallel DNA sequencing and synthesis methods are increasing the speed of DNA synthesis, and there are a number of instruments on the market with the ability to produce hundreds or thousands of DNA sequences.\textsuperscript{10} Many of these systems use 96- or 384-well plates, in which a different DNA sequence is synthesized in each well. Moreover, several manufacturers provide instruments that allow multiple 384-well plates

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Price_Per_Base_of_Synthetic_DNA.png}
\caption{DNA synthesis cost over time.\textsuperscript{9}}
\end{figure}

\textsuperscript{9} Ibid.
to be run on a single instrument, dramatically increasing synthesis throughput.\textsuperscript{11} Research by Rob Carlson, shown in figure 3, indicates that there has been exponential growth in the synthesis community’s ability to create synthetic oligonucleotides and gene-length DNA, which Carlson defines as writing DNA.\textsuperscript{12} Carlson’s data for writing DNA stops in 2008 because no new synthesizing technologies had been released publicly since then. However, he noted that at least two firms—Cambrian Genomics and Synthetic Genomics, Inc. (SGI)—were developing new synthesis and assembly technologies in 2013. In April 2015, one of those companies—SGI-DNA, a subsidiary of SGI—introduced the BioXp\textsuperscript{TM} 3200 System, the first benchtop synthesizer capable of creating genes.\textsuperscript{13} Before the BioXp\textsuperscript{TM} 3200 System was invented, benchtop synthesizers could only produce oligonucleotides.

\footnotesize

\textsuperscript{12} Carlson, “Time for New DNA.”

While there is a high degree of automation in the gene synthesis industry, synthesizing larger genes and whole genomes is more difficult than synthesizing oligonucleotides and short genes because longer stretches of DNA introduce more opportunities for the DNA to recombine or behave in unpredictable ways. As of 2013, the largest sequence that gene synthesis companies reported synthesizing was

---

14 Ibid.
15 Marsic et al., “PCR-Based Gene Synthesis.”
approximately 250 kb. Gene synthesis companies do not typically produce larger genes because there is currently little demand for genes of this size. While entire genomes for certain microorganisms have already been synthesized, these syntheses were performed in academic laboratories for research purposes, not by gene synthesis companies. Moreover, much of the work of creating these larger genomes is performed in a step-wise process that uses smaller stretches of DNA to build ever-larger ones. For example, when researchers at the J. Craig Venter Institute (JCVI) synthesized the large, 1.08 million bp Mycoplasma mycoides genome in 2010, they asked a gene synthesis firm to produce 1,078 gene fragments that were 1,080 bp in length. JCVI researchers then combined the 1,078 fragments into 100 gene assemblies that were 10 kb (10,000 bp) in length. These 100 gene assemblies were then combined into 11 gene assemblies that were 100 kb (100,000 bp) fragments. Finally, they combined the 11 gene assemblies to produce the entire M. mycoides genome.

As mentioned previously, synthetic genomics serves as an enabling technology for synthetic biology because synthetic biologists typically use synthetic DNA for their projects, rather than DNA that has been derived from a host organism. Synthetic biology is an interdisciplinary technology that combines the laboratory techniques involved in genetic engineering and molecular biology with chemistry, engineering, and computer science to design custom biological systems. The ability to redesign or create organisms from scratch could potentially allow scientists to increase their efficiency and their

---

16 One kilobase is equivalent to 1,000 base pairs (bp) of DNA. Therefore, a 250 kb gene is 250,000 bp in length, which is considered a large gene.

understanding of the complex functions of biological systems. While traditional genetic engineering involves transferring existing DNA that is cloned from natural sources, synthetic biology aims to generate and assemble a new genome from chemical—not natural—starting materials.\textsuperscript{18} Synthetic biology simplifies the process of genetic engineering by allowing researchers to produce or purchase sets of genes or entire genomes instead of extracting genes, one at a time, from one species and transferring them into another species using traditional genetic engineering methods.

Researcher University of Exeter Maureen O’Malley and colleagues describe synthetic biology as encompassing three different approaches: genome-driven cell engineering, protocell\textsuperscript{19} creation, and DNA-based device\textsuperscript{20} construction.\textsuperscript{21} These three approaches are exemplified, respectively, in the work of the JCVI, geneticist George M. Church and colleagues at Harvard University, and synthetic biologist Drew Endy of Stanford University. Among other projects, the JCVI researchers are attempting to use synthetic DNA to create a minimal genome. Organisms only use a fraction of the genes in their genome at a given time, and minimal genome efforts, such as the JCVI’s work, seek to create organisms containing only those genes that allow it to survive in a particular environment. After researchers identify this minimal set of genes, they can selectively add other genes of interest, such as those that code for antibiotics or fuels, to the

\textsuperscript{18} Tucker and Zilinskas, “Promise and Perils”; and Czar et al., “Gene Synthesis Demystified.”
\textsuperscript{19} A protocell is a collection of phospholipids that serve as a step towards forming a living cell.
\textsuperscript{20} An example of a DNA-based device is a nanopore or nanochannel whose inner surface has been modified with DNA. Researchers have used these nanopores and nanochannels to study the transportation mechanism and behavior of ions within cells. For more detail, see Dongsheng Liu, Enjun Cheng, and Zhongqiang Yang, “DNA-based Switchable Devices and Materials,” \textit{NPG Asia Materials} 3, December (2011): 109-114, http://www.nature.com/am/journal/v3/n12/full/am2011188a.html.
genome. In this way, the minimal genome serves as a platform, or chassis, that can be augmented with whatever genes a synthetic biologist desires. Constructing minimal genomes allows scientists to determine which genes in an organism’s genome are essential for its survival. Minimal genomes also reduce genetic complexity in an organism, which can make engineered organisms more predictable. The JCVI’s assembly of the 1.08 Mb *M. mycoides* genome, the subject of a case study in chapter 5 of this dissertation, is the group’s first step toward building a minimal genome.

The JCVI’s minimal genome research is part of a larger effort to create a minimal cell that contains the minimum number of components needed to be considered alive or to display some of the functions of a living cell, such as cell division. The purpose of minimal cell research is to produce a well-characterized and predictable chassis into which genes or genomes—perhaps minimal genomes—can be inserted. Minimal cells could serve as platforms that allow researchers to study the function of every gene in a chromosome. This could increase researchers’ understanding of how cells work and help them design synthetic cells that have useful properties, like the ability to produce fuels or pharmaceuticals. Minimal cells differ from minimal genomes because they seek to define the minimal machinery necessary for cells to carry out cellular functions, whereas minimal genomes seek to narrow an organism’s genetic code to only those genes that enable the organism to live in a specific environment.

---


Minimal genome efforts, such as the JCVI’s work, complement protocell creation. For example, one of the ways that researchers can construct a protocell is by reducing the genome in an existing cell so that it contains only those genes that are necessary to sustain its existence. Another approach is to synthesize the basic molecular components of a cell—those components that allow it to self-replicate—and then encapsulate these components into liposomes to create a cell that is semiartificial.\textsuperscript{25} George Church’s effort to engineer biosafety features into a microbial chassis is an example of protocell creation.\textsuperscript{26} In early 2015, Church and colleagues announced that they had reengineered the genetic code of the bacterium \textit{Escherichia coli} so that it would require a synthetic amino acid—one not found in nature—to live.\textsuperscript{27} The purpose of the engineered \textit{E. coli} strain is to serve as a platform, or chassis, for synthetic biology research. Synthetic biologists who use this \textit{E. coli} strain as a platform for their research would not have to worry if the strain escaped the laboratory because it would die without the synthetic amino acid. This massive reengineering also has another advantage: the engineered \textit{E. coli} strain is resistant to viral infections because the viruses cannot recognize the \textit{E. coli}’s genetic code.\textsuperscript{28} This viral resistance has the benefit of protecting the cells from contamination when they are grown in a laboratory. This approach differs from the

\textsuperscript{26} O’Malley and colleagues note that George Church also conducts research on DNA-based devices, but that his work on chassis best fits in the protocol creation category.
JCVI’s minimal genome research because it seeks to radically change the properties of a genome.29

Drew Endy, a bioengineer with a background in civil and environmental engineering, is attempting to make biology easier to engineer by standardizing the parts—the DNA sequences—that synthetic biologists use to design and build the biological systems that are incorporated into cells. To this end, Endy cofounded BioBricks, a registry of standardized and interchangeable biological parts that yield predictable results when inserted into cells or genomes. Founded in 2005, BioBricks provides synthetic biologists a list of over 20,000 parts whose functions have been characterized and validated and can be mixed and matched.30 Drawing on his engineering background, Endy’s approach is to convert synthetic biology to an engineering discipline by developing predictable, modular tools and systems that other synthetic biologists, such as George Church and JCVI researchers, can use as a starting point for building genomes and cells.

Though not mentioned by O’Malley and colleagues, another influential figure in synthetic biology is Jay Keasling, a bioengineer at the University of California, Berkeley. Keasling’s efforts focus on identifying practical applications for synthetic biology. For example, in 2006, Keasling and colleagues engineered the yeast *Saccharomyces cerevisiae* to produce artemisinic acid, a precursor to artemisinin, considered one of the

---

best antimalarial treatments available.\textsuperscript{31} Previously, the only source of artemisinin was the sweet wormwood plant. The supply and cost of the plant is unpredictable, making it difficult to treat the hundreds of millions of malaria cases that arise worldwide each year.\textsuperscript{32} In 2013, pharmaceutical firm Sanofi announced it would begin producing a semisynthetic version\textsuperscript{33} of the drug, and in 2014, the company announced that it had delivered the first large-scale batches of the treatments to parts of Africa.\textsuperscript{34}

Synthetic biologists can use both minimal cells and protocells to study gene function; insert minimal genomes; and add desired genes, such as those that express medically useful compounds. These efforts are further aided by the BioBricks registry, which provides knowledge of the parts, systems, and processes required to devise minimal cells and protocells. The goal of these technologies is to develop cells that can be engineered to produce useful products for peaceful purposes, such as pharmaceuticals and fuels; however, many argue that these same technologies could be used to create dangerous pathogens that could harm humans or the environment.


\textsuperscript{33} The drug is considered semisynthetic because it uses artemisinin from either sweet wormwood or the engineered yeast. The artemisinin is then chemically modified to produce an active form of the drug, such as artesunate, and then combined with another antimalarial drug to lessen the chance that the malaria parasite will develop resistance to artemisinin.

Discourse on Biotechnological Advances and National Security

National security analysts are concerned that terrorists or other malefactors could use synthetic biology for nefarious purposes. Specifically, they are concerned that malefactors could exploit synthetic biology to create select agents, eradicating diseases, or novel pathogens that might have devastating effects on a human or animal population that has not previously encountered them. Security analysts have expressed similar concerns about recombinant DNA technology; however, these traditional methods require access to naturally occurring organisms or their nucleic acids, which analysts considered a deterrent. In contrast, the availability of synthetic DNA, as well as publicly available knowledge of how genes function in a pathogen, has heightened analysts’ concern that malefactors may more easily create highly pathogenic microorganisms de novo, without the need for template DNA. As proof that synthetic genomics increase the threat of biological terrorism, threat assessments often cite the decreasing costs of synthetic genes as well as the widespread availability of necessary materials, including reagents,

---

35 Select agents are pathogens, biological toxins, and associated nucleic acids that pose a potential threat to public health and safety, plant and animal health, and plant or animal production.


protocols, information about DNA sequences, and automated DNA synthesizers. Security analysts surmise that these technological advances provide malefactors an easier, cost-effective way to produce pathogens by either synthesizing their own genes or purchasing genes from commercial firms.

Several experiments in the past decade have exacerbated concerns that gene synthesis and synthetic biology could be used for nefarious purposes. In 2000, Australian researchers modified the mousepox virus using synthesized genes, unintentionally rendering it virulent in mice that had previously been vaccinated against it. In 2002, researchers at the State University of New York at Stonybrook detailed the de novo synthesis of poliovirus from commercially purchased oligonucleotides. The research showed that viruses could be synthesized in the absence of previously infected cells so long as researchers knew the sequence of the pathogen’s genome. In 2003, JCVI researchers published an alternative, quicker method for synthesizing the genome of a larger infectious microorganism, the phiX bacteriophage. In 2005, government researchers synthesized the influenza virus that was responsible for the Spanish Flu pandemic of 1918 and 1919, which caused 30 to 50 million deaths worldwide. A 2006 article in The Guardian described how a reporter had ordered part of the smallpox genome from a United Kingdom gene synthesis firm for $40. The article suggested that it

38 Carlson, “Pace and Proliferation”; Garrett, “Biology’s Brave New World”; Randerson, “Revealed: The Lax Laws” (see chap. 1, n. 7); and Tucker and Zilinskas, “Promise and Perils.”
40 Cello, Paul, and Wimmer, “Chemical Synthesis of Poliovirus” (see chap. 1, n. 5).
42 Tumpey et al., “Double-Edged DNA” (see chap. 1, n. 6).
would be similarly easy for terrorists to obtain the “basic ingredients of biological weapons.”

In 2008, scientists in Tennessee synthesized a bat severe acute respiratory syndrome (SARS)-like coronavirus and converted it into an infectious clone. And in 2011, university teams in the Netherlands and the United States enhanced the transmissibility of the H5N1 strain of bird flu virus so that it could be transmitted among ferrets; prior to that experiment, the virus had not had the capacity for airborne transmission among mammals.

Threat assessments generally point out that although biotechnologies have yielded revolutionary benefits for medicine and the environment, they also escalate the likelihood of bioterrorism or biowarfare due to rapidly advancing technological capabilities, expanding adoption by scientists worldwide, and increasingly sophisticated instruments and protocols that decrease the skills needed to use the technologies. The assessments often comingle information about biotechnology’s benefits with dire predictions of the future. For example, Stanford researcher Christopher Chyba asserts that although biotechnological advances will one day enhance human health, they nevertheless pose a “grave challenge” to arms control, due to their exponential growth and rate of proliferation around the world, their potential to yield agents that could reproduce in the natural world, and their increasing availability and utility for small groups or

---

43 Randerson, “Revealed: The Lax Laws.”
individuals. Biologists Ronald Atlas of University of Louisville and Malcolm Dando of the University of Bradford believe as more knowledge of life sciences and biotechnology is used to improve human health, the risk of its misuse increases. This misuse, the authors contend, will likely transform the life sciences into the “death sciences.” University of California, Davis, microbiologist Mark Wheelis asserts that because biotechnology and bioinformatics have increased researchers’ understanding of living organisms’ physiology—and because this knowledge has proliferated quickly around the world—there is an increased likelihood that the knowledge will be used for nefarious purposes.

Many of these concerns about the global proliferation of biotechnology, and the pace thereof, are parroted in U.S. government assessments. The 2008 World at Risk Report characterizes advances in the life sciences as a “global revolution,” while the National Strategy for Countering Biological Threats contends that advances in the life sciences are progressing globally at an “unprecedented rate,” suggesting that this proliferation automatically increases the risk of a biological attack. Reporting that advances in gene synthesis created possibilities that were “both grand and terrifying,” a 2012 article in The Atlantic claimed that “exponential advances” in biotechnology have

---

48 Wheelis, “New Biology New Weapons?”
greatly diminished previous problems working with DNA. A 2015 *Atlantic* article by Patrick Tucker echoed these concerns, stating that the “quickening pace of genetics research has plenty of scientists worried.”

Assessments also discuss biotechnology’s potential for increasing the lethality of biological agents, including the ability to make novel pathogens. Chyba and Alex Greninger, for example, maintain that biotechnology will “inevitably place greater destructive power in the hands of smaller groups of the technically competent.”

Meanwhile, Wheelis contends that some biotechnological advances could enable the creation of new weapons that might “transform the nature of combat in unprecedented ways.” He goes on to predict that within the next two decades or more, the world can expect the further development of pathogens that are increasingly contagious and environmentally stable; that evade diagnosis and treatment; or that disable victims permanently, among other effects. He further argues that “it will not be long before completely novel synthetic viruses are produced.” Defense Intelligence Agency analysts James Petro, Theodore Plasse, and Jack McNulty contend the threat posed by genetically modified traditional agents will plateau because only a finite number of genetic modifications can be used to enhance traditional agents without altering them beyond

---


53 Wheelis, “New Biology New Weapons?”

54 Ibid.

55 Ibid.
recognition of the parental strain. The authors believe that, unlike the threats posed by traditional and genetically modified agents, the threat posed by novel biological warfare agents will continue to expand indefinitely in parallel with advances in biotechnology.

These concerns about biotechnologies are echoed in threat assessments that are specific to synthetic biology and its enabling biotechnologies. These assessments highlight the concern that states or terrorist organisms might exploit synthetic biology—and, by extension, synthetic genomics—for nefarious purposes, such as by creating novel pathogens or traditional biological agents or biological weapons. A 2013 article that appeared in *Foreign Affairs* claimed, “All the key barriers to the artificial synthesis of viruses and bacteria have been overcome, at least on a proof-of-principle basis.” A recent Council on Foreign Relations report asserted that the deliberate misuse of synthetic biology is the greatest risk to U.S. national security, a claim that was echoed in a *Global Challenges Foundation* report describing global risks that pose a threat to human civilization and, possibly, all human life. According to the report, one of most impactful risks of synthetic biology is the creation of an engineered agent that targets humans or the environment. The report concluded, however, that this scenario was unlikely to result in global extinction.


57 The authors define advanced biological warfare agents as novel biological warfare agents created using biotechnological applications.


59 Garrett, “Biology’s Brave New World.”


61 Ibid.
One of the key weaknesses of these assessments is that they equate accessibility of material and technologies to capability, therefore concluding that terrorists or other malevolent actors could create dangerous agents. Missing in these reports is a rigorous analysis of the type of skills that are required to carry out such experiments and which factors might enable or prevent successful outcomes. STS researchers Claire Marris, Catherine Jefferson, and Filippa Lentzos reviewed academic and nonacademic literature published since 1999 that mention synthetic biology and biosecurity and found that the literature does not draw upon or cite STS research.62 Indeed, the phrase “tacit knowledge” did not appear in any of the reports surveyed. Most articles that do mention the skills required to use biotechnologies do so in vague terms; for example, by noting that the technologies can be exploited by those with “technical ability” or by saying that the experiments are “not the sort of thing a terrorist . . . could easily attempt in a kitchen.”63 One assessment claimed that unspecified “scientific and medical know-how” was all that was needed for a member of the Islamic State of Iraq and the Levant (ISIL) or other terrorist groups to deliberately misuse synthetic biology.64 Another assessment explained that malefactors with undefined “scientific or medical know-how” could join terrorist organizations and carry out biological attacks, or that malefactors could use do-

it-yourself (DIY) biology laboratories to create pathogens like Ebola or influenza. Yet another assessment claimed that due to exponential growth in technologies such as DNA sequencing and synthesis, as well as falling costs, genetic engineering could “easily” be used for destructive purposes. A 2015 article in the popular journal *The Atlantic*—alarmingly entitled “The Next Manhattan Project”—contended that there is “wide agreement” that the barriers to entry for creating new biological organisms, such as those that could kill millions of people, are decreasing.

The few studies that do attempt to characterize skill in greater detail do so by describing the ease with which scientists can use associated materials and follow protocols. One assessment explained that commercially available reagent kits and associated protocols decrease the skill required to produce synthetic genes and that using these kits and their “simple recipes” is “more complicated than baking cookies, but . . . less complicated than making wine and beer.”

The threat assessment discourse on biotechnologies highlight three consistent themes from national security analysts: (1) a belief that technology has its own momentum that is not subject to human shaping, (2) an assumption that technology is inevitably used for nefarious purposes, and (3) a reliance on technology forecasting to predict how biotechnologies can be used in the future. In the following sections, I deconstruct these three assumptions.

---

65 Gronvall, “Mitigating the Risks.”
66 Hessel, Goodman, and Kotler, “Hacking the President’s DNA.”
68 Carlson, “Pace and Proliferation,” 209.
**Technological Determinism**

Much of the discourse about biotechnology’s influence on national security rests on a deterministic view of how technology is developed and used. In technology determinism, new technology developments are assumed to follow an existing, typically linear trajectory, with the technology having an internal momentum separate from any human influence or control. Furthermore, this view primarily focuses on the developments of necessary materials and information, rather than considering the contexts in which new technologies are developed or used. Consequently, the deterministic view does not consider the social or environmental factors—not to mention the potential for human error—which may be introduced into a technology’s development or adoption, thus delaying or stalling a presumed linear trajectory of advancements.

As noted in the previous section, threat assessments typically address the variable of human interaction with biotechnologies by noting that the technologies must be used by individuals with technical ability, a vague description that cannot capture how individuals use biotechnologies in meaningful terms. Threat assessment authors state that the use of biotechnologies requires competent individuals—defined as those with “technical expertise” or a “knack for science”—but they do not investigate what constitutes this necessary expertise; the assessments are nevertheless published and disseminated without this crucial information.69

---

69 Wheelis, “New Biology New Weapons?”; and Hessel, Goodman, and Kotler, “Hacking the President’s DNA.”
Other discussions of human involvement in biotechnology focus on the increasing automation of associated instrumentation and the availability of kits that contain ready-made reagents and easy-to-follow experimental protocols—or “simple recipes,” as one analyst described them. Analysts argue that these factors eliminate or significantly reduce the need for human involvement as well as any potential errors that a human would introduce. However, those who adhere to this view fail to recognize that automation does not necessarily remove all need for skill. In fact, in some cases, automation produces the need for new types of skill, such as knowing how to use the equipment and solving problems when the equipment malfunctions.

New technological advances and automated equipment can also lead to the technology being used in unexpected ways or even abandoned, options that are not anticipated by technological enthusiasts. For example, research on the automobile industry’s early days shows that automobiles, like many technologies, are characterized by interpretive flexibility; that is, different social groups associate different meanings with the automobile. Historian Ron Kline and STS researcher Trevor Pinch’s research shows that automobile manufacturers had great influence on the form that the automobile took initially but were unable to control how automobiles were used once customers bought them. Manufacturers initially marketed automobiles as a form of transportation for city drivers, but rural owners adapted the automobile to suit farm life and help them perform work, rather than just provide a means of transportation. In applying this

70 Carlson, “Pace and Proliferation,” 209.
72 Ibid.
example to biotechnology, interpretive flexibility suggests that users may not use synthetic biology equipment in the ways that are most feared, for example, as a tool for conducting terrorist acts.

In addition to paying inadequate attention to the social context in which biotechnology is developed and used, technological determinism is also evident in reports that equate an individual’s or a group’s intent to use biotechnology for nefarious purposes with its ability to do so. In a 2010 article assessing terrorist acquisition of weapons of mass destruction (WMD), Rolf Mowatt-Larsen, Senior Fellow at Harvard’s Belfer Center for Science and International Affairs and former Central Intelligence Agency officer, points to the clear intent of al-Qaeda and similar groups to acquire and use WMD.\textsuperscript{73} In the article, Mowatt-Larsen asks why, despite the predictions in a June 2003 U.S. government warning, the United States has not experienced a WMD attack. At no point in the article was the difficulty associated with developing WMD addressed or offered as a possible reason for why a WMD attack has yet to occur.\textsuperscript{74} Similarly, in a 2010 statement by the Commission on the Prevention of Weapons of Mass Destruction Proliferation and Terrorism that gave the United States poor ratings for addressing WMD proliferation and terrorism, Senator Bob Graham stated that the United States no longer has the “luxury of a slow learning curve, when we know al-Qaeda is interested in

\textsuperscript{73} Rolf Mowatt-Larsen, “Al Qaeda Weapons of Mass Destruction Threat: Hype or Reality?,” (paper, Belfer Center for Science and International Affairs, January 25, 2010).

\textsuperscript{74} Ibid.
Here again, there is no detailed discussion about the skill level that would be needed by al Qaeda to develop such weapons.

All of these pronouncements assume that individuals or organizations have the capability to create a biological agent or weapon so long as they have the intent to do so and that their capability to create an agent or weapon is synonymous with possessing materials and information, which are seen as ubiquitous. This simplistic conclusion defies the commonly accepted definition of “threat,” which factors in both an adversary’s intent to commit an act and its capability to do so. As highlighted by security researcher Sonia Ben Ouagrham-Gormley and STS researcher Kathleen Vogel, security analysts who adopt this assumption assert that individuals or organizations are capable of creating a biological agent or weapon so long as they have access to materials and information—with the variable of human expertise remaining nebulous and unexplored. Capability involves more than just access to technologies, materials, and information. Thus, threat assessments should also address the skill sets, expertise, and other social factors that are necessary to use technologies, materials, and information.

**Inevitability of Nefarious Use**

The assumption that biotechnology will inevitably be used for nefarious purposes is a mainstay of national security assessments. The 2009 *World at Risk* report asserts that

---


76 “DHS Lexicon,” *Department of Homeland Security*, September 2008, accessed March 7, 2015, https://www.dhs.gov/xlibrary/assets/dhs_risk_lexicon.pdf. The DHS defines capability as an adversary’s means to accomplish a mission, function, or objective, and notes that adversary capability is considered along with adversary intent when estimating the likelihood of a terrorist attack. The DHS defines intent as the determination to achieve an objective.

77 Ben Ouagrham-Gormley, *Barriers to Bioweapons* (see chap. 1, n. 11); and Vogel, *Phantom Menace* (see chap. 1, n. 11).
rapid innovations in the biological sciences and the malevolent intentions of proliferators will “sooner or later” intersect. Chyba and Greninger contend that basic medical research into the human immune system and the human genome, in addition to experiments with pathogens, will inevitably “point the way towards increasingly powerful . . . methods to cause human disease.” In many cases, the inevitability is associated with an actor’s intent to acquire or use WMD. This assumption is also parroted in popular media. For example, a 2012 article published in The Atlantic—“Hacking the President’s DNA”—stated, “Whenever novel technologies enter the market, illegitimate uses quickly follow legitimate ones. A black market soon appears.” Thus, according to the article, criminals and terrorists will “surely soon turn to synthetic biology.”

Many security assessments also imply that malefactors’ misuse of biotechnology advancements is inevitable by framing their discussions primarily on the possibilities of hostile use. For example, Wheelis stated in 2004 that the biotechnology revolution is making it possible for biology to become a “full-fledged military technology.” In a 2005 report, Wheelis and Dando suggest that the biological, medical, and legal communities face the “near certainty” that biology will become the next major military technology because “no major technology with military utility has not been exploited for

---

78 Graham et al., World at Risk, 12–13.
79 Chyba and Greninger, “Biotechnology and Bioterrorism, 150.
81 Hessel, Goodman, and Kotler, “Hacking the President’s DNA.”
82 Wheelis, “New Biology New Weapons?”
As such, the “wholesale militarization of biology will be an integral part” of the revolution in biotechnology. Although these assessments do discuss biotechnology’s positive role in advancing healthcare and the environment, as well as countering the effects of biological weapons, their emphasis on the hostile potential implies that biotechnology’s risks might outweigh its benefits. In contrast, medical, pharmaceutical, energy, and biotechnology industry reports typically focus on advances in biotechnology that can be used for practical, peaceful purposes. Moreover, the contention that biotechnology might become a military technology runs counter to a fact that many researchers have asserted: biological agents do not make good battlefield weapons, due to their unpredictability.

Figure 4 provides a visual depiction of this hostile frame in threat assessments on synthetic genomics and synthetic biology by presenting text from the assessments in a word cloud. In the figure, the words that are used most often appear largest, while the least-used words are in smaller text. To generate the graphic, I extracted statements about synthetic genomics and synthetic biology from twelve randomly selected threat assessments, eliminated common words (for example, “the,” “very,” “these,” etc.), and I standardized related terms (for example, changed “bioweapon” to “bioweapons” and

84 Ibid., 567.
“designing” to “design”). The graphic shows the preponderance of language that relates to the technologies’ hostile uses (for example, “risks,” “biowarfare,” and “misuse”). It also shows some words that relate to the technologies’ peaceful uses (for example, “countermeasures,” “legitimate,” and “beneficial”). The figure demonstrates that security assessments more often refer to the technologies’ hostile applications, rather than their peaceful applications, and there is little mention of the skills required to develop these weapons. Since it is national security analysts’ job to focus on hostile implications of the technologies, one could argue that it is not surprising that they place more focus on the hostile uses of those technologies, not the peaceful uses. However, as discussed in the previous sections, the national security analysts often fail to consider the expertise required to carry out these hostile intentions.
Given the discourse about biotechnologies in many threat assessments, it is easy to understand why advances in the technologies would elicit alarmist and suspicious responses from politicians and the public. Researcher Jessica Stern of Harvard University believes these responses are based on the perception that, due to innate characteristics such as contagiousness, biological weapons are dreaded risks. As dread risks, biological weapons elicit a disproportionate level of fear and horror. Because biotechnologies are thought to enable the creation of biological weapons, it follows that they themselves are dread risks. However, STS literature suggests that using biotechnologies is more difficult.

---

than most national security analysts perceive it to be, which ultimately suggests that access to biotechnologies may not increase the likelihood of biological terrorism or warfare after all.

**Theoretical Possibilities**

Much of the national security discourse is also based on the use of agents and scenarios that are only theoretically possible. A 2003 article by Defense Intelligence Agency analysts Petro, Plasse, and McNulty predicted that advanced biological warfare agents will be able to target specific physiological systems, such as the cardiovascular, immunological, neurological, and gastrointestinal systems, and that bioengineers of the future will be able to engineer agents that target specific biological processes.\(^{87}\) Similarly, Wheelis predicted that, in the long term, adversaries would be able to develop synthetic viruses and prions, stealth pathogens that are engineered to become latent after a period of replication, ethnic-specific human pathogens, and pathogens that cause ethnic-specific autoimmune diseases and may affect fertility.\(^{88}\)

The majority of threat assessment predictions are not accompanied by references to current intelligence reports or other official sources, indicating that they are based on speculation. In essence, they are technology forecasts, which are notoriously unreliable and vague. In 2012, an analytical consulting firm based in Alexandria, Virginia, conducted a retrospective analysis of 1,000 technology forecasts for a U.S. Department of Defense office charged with maintaining the U.S. military’s technological edge. The

---

\(^{87}\) Petro, Plasse, and McNulty, “Biotechnology: Impact.”

\(^{88}\) Wheelis, “New Biology New Weapons?” It is important to note that several analyses have shown that although the South African biological weapons program attempted to develop ethnic-specific weapons, they were unsuccessful. See, for example, Ben Ouagrham-Gormley, *Barriers to Bioweapons.*
office routinely received technology forecasts from the intelligence community, think
tanks, working groups, and other sources and commissioned the study to determine if
factors such as the institutional or geographic origin of forecasts or forecast methodology
influenced accuracy. The analysis revealed that forecasts that are based on mathematical
and statistical techniques and those that predict shorter time horizons—the length of time
between when a forecast was made and the predicted data an event will occur—are more
accurate than all other forecasts, including those made by individuals or groups who were
considered experts in their fields. The researchers also encountered a large number of
technology forecasts that were too vague to be evaluated; that is, they lacked sufficient
information about when an event should occur or what methodology the authors used to
generate the forecast. This indicates a key failure in the process to develop meaningful
technology assessments for decision makers. Together, these findings demonstrate that
many current threat assessments are unreliable and too vague to allow the end-users to
make well-informed decisions. Thus, without credible information or sources to
substantiate their predictions, these speculative threat assessments may only result in the
United States engaging in what one analyst described as an “arms race with ourselves.”

The Cold War and post-Cold War periods are replete with examples of how U.S.
technological forecasting led to worst-case security scenarios that were far from reality
and led to poor policymaking. One example is the findings of the Iraq Survey Group, an
international team whose mission was to identify signs of WMD production in Iraq

89 Frank von Hippel, “Must We Have a BW Arms Race with Ourselves?” Final Session, Princeton
University Seminar Series: The Biodefense Challenge: How Should the Life-Science Research Community
workshop/FvH-BWArmsRace040521.pdf.
following the U.S. invasion of that country in 2003. Prior to the U.S. invasion, the intelligence community reported that Iraq was developing biological weapons. These intelligence reports were key factors in the American decision to invade Iraq. The Iraq Survey Group conducted an extensive investigation while in Iraq, including debriefing former members of the Iraqi regime, conducting technical analyses of materials collected in the country that were suspected of being related to its WMD program, and collecting human intelligence about the WMD program. The group found no evidence that Iraq had been developing biological weapons; there was no equipment suitable for the type of biological weapons capability that the intelligence community described, no production sites, no documents pertaining to the program, and no information from interviews with individuals suspected of being involved in the program. The group documented these findings in a 2004 report, commonly referred to as the Duelfer Report.

In addition, the Report of the Commission on the Intelligence Capabilities of the United States regarding Weapons of Mass Destruction to the President of the United States, submitted March 31, 2005, examined intelligence community practices that led to the false assessments about Iraq’s supposed WMD program. The report stated that the intelligence community was “dead wrong in almost all of its pre-war judgments about Iraq’s weapons of mass destruction” and that the intelligence about Iraq’s WMD

---

91 Iraq did develop biological weapons prior to the U.S. invasion in 2003. However, United Nations Special Commission shut down Iraq’s program after the first Gulf War.
programs “was a major failure.” It went on to explain that the principal cause of the intelligence failure was analytical: the intelligence community had collected too little information about Iraq’s WMD program, and much of the information they did collect was “worthless or misleading.” In addition, analysts had biased their assessments of Iraq’s WMD capability by making assumptions about Saddam’s intentions; in other words, they paid inadequate attention to Saddam’s capabilities. Additionally, they did not communicate the confidence they had in their sources or their assessments.

The United States has adopted a number of policies and other tools to reduce the likelihood that terrorists and other malefactors will be able to develop bioweapons and wage biowarfare using dual-use technologies. These policies and tools are explained in the next section.

**Threat-Mitigation Policies: Restricting Access to Dangerous Pathogens and Associated Genetic Material**

A number of U.S. policies are designed to mitigate the perceived threats stemming from advances in biotechnologies such as synthetic genomics and synthetic biology. These policies, which focus on restricting access to pathogens and their genetic material, include the Select Agent Rule (SAR), the Australia Group, Export Administration Regulations (EAR), the Biological and Toxin Weapons Convention (BTWC), and the Chemical Weapons Convention (CWC).

---

94 Ibid., 3.
95 For more information about failures with intelligence analysis, see Vogel, *Phantom Menace.*
The Select Agent Rule

The SAR was instituted in 1997, after an individual affiliated with a domestic hate group illegally obtained *Yersinia pestis*—the causative agent of the plague—by mail in 1996. In response to this incident, Congress passed the Antiterrorism and Effective Death Penalty Act of 1996 (Public Law 104-132, Section 511), which required the U.S. Department of Health and Human Services (DHHS) to regulate the transfer of select agents like *Y. pestis*. In 1997, the DHHS published the SAR (Additional Requirements for Facilities Transferring or Receiving Select Agents, Title 42 Code of Federal Regulations [C.F.R.], Part 72), which regulated the access, use, and transfer of select agents by requiring anyone shipping or receiving select agents to register with the DHHS. However, the law only regulated the transfer and acquisition of select agents, not their possession; updates to the SAR have since been passed to help cover that gap.

In 2001, the United States passed the Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act (USA PATRIOT Act), which prohibits “restricted persons” from shipping, possessing, or receiving select agents and toxins and criminalizes possession of select agents in quantities that cannot be justified for protective, research, or other legitimate purposes. Under the PATRIOT Act, restricted persons are prohibited from shipping, receiving, possessing, and transporting select agents.

---

96 The DHHS issued a final rule in 2008 that rescinded 42 C.F.R. 72 and replaced its requirements for handling select biological agents and toxins with the regulations that are now found in 42 C.F.R. 73.
97 Additional Requirements for Facilities Transferring or Receiving Select Agents, 42 C.F.R. Part 72.
98 “The USA Patriot Act: Preserving Life and Liberty,” *Department of Justice*, accessed November 6, 2015, http://www.justice.gov/archive/ll/highlights.htm. “Restricted persons” include citizens of countries on the U.S. State Department’s list of state sponsors of terrorism, individuals with a history of drug abuse or mental instability, or individuals with a criminal background.
The Public Health Security and Bioterrorism Preparedness and Response Act (Public Law 107-188) was passed the following year. It mandates that individuals who possess, use, or transfer select agents or toxins that are considered a threat to public, animal, or plant health must register with the Centers for Disease Control and Prevention (CDC) or U.S. Department of Agriculture (USDA) and undergo a Federal Bureau of Investigation (FBI) background check to ensure the individual does not pose a security threat. The law also mandates that laboratories working with select agents and toxins implement safety and security measures to ensure the agents and toxins do not escape the laboratories. 99

In order to meet the requirements of the Public Health Security and Bioterrorism Preparedness and Response Act, the DHHS’s CDC developed rules on the possession, use, and transfer of select agents and toxins, while the USDA’s Animal and Plant Health Inspection Service (APHIS) developed rules on the possession, use, and transfer of biological agents and toxins. These efforts culminated in three lists—effective as of 2005—that comprise the SAR: the Possession, Use, and Transfer of Select Agents and Toxins (Title 42 C.F.R., Part 73) and the Possession, Use, and Transfer of Biological Agents and Toxins (Title 7 C.F.R., Part 331, and Title 9 C.F.R., Part 121). 100

The CDC regulates some of the select agents while the APHIS regulates others. Some overlap agents, such as *Bacillus anthracis*, appear on more than one list and are regulated by either the CDC or APHIS. The SAR accounts for (1) nucleic acids that can

---

provide infectious forms of any of the select agent viruses, (2) recombinant and/or synthetic nucleic acids that encode for the functional forms of select toxins if the nucleic acids can be expressed in vivo, in vitro, or are in a vector or recombinant host genome and can be expressed in vivo or in vitro, and (3) select agent and toxins that have been genetically modified. The SAR does not address the use, possession, or transfer of synthetic gene fragments that by themselves are unable to produce a functional form of a select agent or toxin, such as oligonucleotides and synthetic genes.

In 2010, President Barack Obama signed Executive Order 13546, Optimizing the Security of Biological Select Agents and Toxins in the United States, which directed the DHHS and the USDA to consider reducing the number of agents and toxins on the select agent list, to prioritize the select agents and toxins by assigning them to different tiers based on their risk of being deliberately misused and their potential to cause mass casualties or devastate the economy or critical infrastructure, and to establish physical and information security standards for select agents with the highest risk of misuse.

Appendix A provides the latest Select Agents and Toxins list.

The Australia Group
Like the SAR, the Australia Group aims to restrict access to highly pathogenic microorganisms and toxins. The Australia Group was formed in 1985 in response to

---

101 For more information, see the website for the Federal Select Agent Program: http://www.selectagents.gov/history.html.


Iraq’s use of chemical weapons in the Iran-Iraq War, which Iraq had obtained legally from Western countries.\textsuperscript{104} The group began with 15 member countries—including the United States—and has since grown to more than 40 members (Australia Group 2007). The Australia Group’s member countries control the export of materials and technologies that are thought to contribute to the proliferation of biological and chemical weapons.

These restricted materials and technologies are articulated in the group’s “common control lists” and include highly pathogenic microorganisms and biological and chemical manufacturing equipment. The common control lists also regulate (1) genetic elements that contain nucleic acid sequences associated with the pathogenicity of any of the microorganisms on the list or that contain nucleic acid sequences coding for any of the toxins on the list, or the toxins’ subunits, and (2) genetically modified organisms that contain nucleic acid sequences associated with the pathogenicity of any of the microorganisms on the list or that code for any of the toxins on the list, or the toxins’ subunits.\textsuperscript{105}

Genetic elements include those that are wholly or partially chemically synthesized. If the synthetic genetic element does not code for part of a controlled agent or toxin, the Australia Group’s guidelines do not apply.\textsuperscript{106} Moreover, if genetic elements are not associated with a genetic sequence that is responsible for pathogenicity, the


elements do not fall under the guidelines, even if they are derived from an organism on
the common control list. However, this is contingent upon researchers knowing the
function of every gene in an organism’s genome, which is not possible for every
organism. It is likely that genetic elements associated with an organism’s pathogenicity
could be exported simply because researchers have not yet learned that the sequence is
associated with pathogenicity. Additionally, the Australia Group has no enforcement
mechanism; its members have no legally binding obligations to each other. Appendix B
shows the materials that are restricted by the Australia Group.

**Export Administration Regulations**

In addition to regulating the export of materials and technologies found in the
Australia Group’s common core lists, the United States also adheres to the U.S.
Department of Commerce’s Export Administration Regulations (EAR), which also
regulate the export of dual-use items. The EAR controls goods and technologies that are
articulated in its Commerce Control List (CCL). Specifically, the EAR directs the U.S.
Food and Drug Administration and Drug Enforcement Administration to control the
export of viruses and bacteria on the Australia Group’s “List of Human and Animal
Pathogens and Toxins for Export Control” (see appendix B), viruses on the select agents
list, toxins identified on the Australia Group’s list, and several fungal species and plant
pathogens.

---

107 Ibid.
108 Bureau of Industry and Security, “Category 1 - Special Materials and Related Equipment, Chemicals,
‘Microorganisms,’ and ‘Toxins’” (Unofficial electronic export administration regulation files created by
Included in the EAR are genetic elements and genetically modified organisms that contain nucleic acid sequences associated with these microorganisms’ pathogenicity or code for the controlled toxins. This means that oligonucleotides or synthetic genes that contain sequences associated with the pathogenicity of microorganisms on the EAR’s CCL must be licensed if they are to be exported. Unlike many other portions of the EAR, regulations on human, animal, and plant pathogens and toxins and their associated genetic material apply regardless of destination; for example, exports to China require the same licensing as exports to Canada or Japan. If a gene synthesis firm received an order for a gene associated with a microorganism or toxin on the CCL, the firm would need to apply for a license to export that gene and wait for the government to complete its review process, which can take up to two months. This delay might preclude specific customers from purchasing DNA from U.S. firms, and instead customers might purchase DNA from firms in countries that do not restrict access to dual-use materials.

**Biological and Toxin Weapons Convention**

In addition to instituting controls to prevent malefactors from acquiring materials, the U.S. government also uses treaties to prevent states from misusing materials and biotechnologies. The 1972 Biological and Toxin Weapons Convention (BTWC) bans member states from developing, producing, stockpiling, or otherwise acquiring and retaining (1) microbial or other biological agents and toxins (regardless of their origin or method of production) of types and in quantities that have no justification for prophylactic, protective, or other peaceful purposes and (2) weapons, equipment, or

---

110 Bar-Yam et al., “Regulation of Synthetic Biology.”
means of delivery designed to use such agents for hostile purposes or armed conflict. The Fourth Review Conference that took place in 1996 reaffirmed that Article I of the convention unequivocally covers all microbial or other biological agents or toxins, naturally or artificially created or altered, as well as their components—whatever their origin or method of production. This latter clause could be applied to synthetic DNA, a component of a synthetic organism that could potentially be used for hostile purposes. A 2011 report that was published before the Seventh Review Conference that year examined the impact of several science and technology trends that are relevant to the BTWC, including synthetic biology and synthetic genomics, supported regulatory and self-regulatory measures such as sequence screening, and urged continued monitoring of advances in synthetic biology and the implications of those advances.\footnote{John Hart and Ralf Trapp, “Science and Technology and their Impacts on the Biological and Toxin Weapons Convention” (a Stockholm International Peace Research Institute synthesis report, December 2011).} While the BTWC addresses the use of synthetic DNA for nefarious purposes, it lacks formal verification or enforcement measures and is therefore difficult to use as a tool for ensuring that member states do not misuse synthetic biology. Moreover, the BTWC applies to nation states, not individual terrorists or terrorist groups.

**Chemical Weapons Convention**

In contrast to the BWTC, the 1993 Chemical Weapons Convention (CWC), which prohibits the development, production, acquisition, stockpiling, retention, transfer, and use of chemical weapons by states parties, has a strong verification regime, but it fails to
address the chemical synthesis of biological molecules. However, the Organization for the Prohibition of Chemical Weapons’ Scientific Advisory Board (SAB) recommended the establishment of a temporary working group in 2011 to explore the increasing convergence of chemistry and biology, as well as to describe implications for the CWC. A June 2014 report produced by the SAB’s temporary working group noted that there is currently no known advantage in trying to produce classical chemical warfare agents through biological means but biological means, like synthetic biology, would be more effective in producing precursors or toxins. The report concluded that with the development of biologically mediated chemical production processes, like synthetic biology, producing scheduled chemicals as warfare agents would be more difficult than obtaining such materials by traditional synthetic chemistry. However, recent research by Stanford University bioengineer Christina Smolke and colleagues demonstrated the possibility of using synthetic biology to create chemical compounds. In 2015, the researchers engineered two yeast strains to convert sugar into thebaine and hydrocodone, two opioid compounds that are used as painkilling medications. Though thebaine and hydrocodone are primarily used for medicinal purposes, they could potentially be used as chemical weapons with lethal effects.

Table 3 summarizes the purposes and limitations of each of the abovementioned regulations, ultimately demonstrating how the current regulatory framework is not sufficient to prevent malefactors from using synthetic genomics to chemically synthesize DNA for highly pathogenic microorganisms.

### Table 3. Summary of U.S. regulatory framework and shortfalls.

<table>
<thead>
<tr>
<th>Regulation</th>
<th>Purpose</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAR</td>
<td>Limits access to specific pathogens, toxins, and associated genetic material to individuals who have legitimate need to work with select agents and who have passed a background check with the FBI.</td>
<td>• Does not regulate the use, possession, or transfer of synthetic genes that by themselves cannot produce a functional form of a select agent or toxin. • Select Agents and Toxins List is not as extensive as Australia Group list.</td>
</tr>
<tr>
<td>Australia Group</td>
<td>Controls the export of materials and technologies that are thought to contribute to the proliferation of biological and chemical weapons. Includes highly pathogenic microorganisms and their associated genetic elements.</td>
<td>• Only applies to the Australia Group’s 41 member countries; nonmembers could export restricted materials. • Lacks an enforcement mechanism; members have no legally binding obligations to one another. • Does not apply to genetic elements that are not associated with an organism’s pathogenicity, even if the organism itself is regulated.</td>
</tr>
<tr>
<td>EAR</td>
<td>Regulates the export of dual-use items, including viruses on the select agents list, toxins on the Australia Group list, and associated genetic elements and genetically modified organisms.</td>
<td>• Receiving a license to export a dual-use gene could take months; might preclude customers from purchasing DNA from U.S. firms</td>
</tr>
<tr>
<td>BTWC</td>
<td>Bans member states from developing, producing, stockpiling, acquiring, or retaining microbial or other biological agents and toxins and their components of types and in quantities that have no justification for prophylactic, protective, or other peaceful purposes; and weapons, equipment, or means of delivery designed to use such agents for hostile purposes.</td>
<td>• Lacks formal verification or enforcement measures. • Applies to nation states, not individuals or organizations.</td>
</tr>
<tr>
<td>Regulation</td>
<td>Purpose</td>
<td>Limitations</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>CWC</td>
<td>Prohibits the development, production, acquisition, stockpiling, retention, transfer, and use of chemical weapons by states parties.</td>
<td>• Does not address the chemical synthesis of biological molecules.</td>
</tr>
</tbody>
</table>

**Proposed New Framework to Address the Potential Threats of DNA Synthesis**

As the previous section shows, the current policies are insufficient to regulate the risks associated with synthetic genomics and synthetic biology. In 2004, the National Science Advisory Board for Biosecurity (NSABB) was established to advise the U.S. government on strategies for preventing the misuse of dual-use research, such as synthetic genomics. In addition to assessing the national security concerns associated with the synthesis of select agents, the NSABB was also charged with identifying, assessing, and recommending strategies to address concerns associated with work conducted in the field of synthetic biology.

In 2006, the NSABB convened a working group to assess whether synthetically derived genes were adequately covered by the regulatory framework for select agents and to determine what strategies and mechanisms could prevent or mitigate the misuse of synthetic genomics.\(^{115}\) Although the NSABB working group noted that synthetic DNA is addressed in the legal framework, it stated that clarification is needed to ensure compliance with existing requirements and that regulatory options need to keep pace with the speed of advances in the field. The working group also found that malefactors could evade the existing regulatory framework by chemically synthesizing select agent DNA or

\(^{115}\) NSABB, “Synthesis of Select Agents” (see chap. 2, n. 37).
obtaining it from commercial companies, and that the technology, reagents, and sequence
information necessary to do so are readily accessible.\textsuperscript{116}

The NSABB recommended that (1) the DHHS and USDA clarify and harmonize
guidance for synthetic DNA providers about which genetic elements and genomes are
covered by Title 42 C.F.R., Parts 73 and 74, and increase awareness among synthetic
DNA providers about the SAR; (2) the U.S. government develop a process that synthesis
companies could use to screen orders for similarities to select agent sequences; and (3)
experts from the scientific community determine if it is feasible to develop a new
regulatory framework based on the predicted features encoded by genes, such as
virulence and pathogenicity, among other recommendations (NSABB 2006).\textsuperscript{117}

Despite all of the recommendations in their report, the NSABB working group did
not describe or evaluate the level or type of expertise required to synthesize DNA. Yet
the pivotal question remains: Would terrorist groups and other malefactors have the skills
required to create select agents from synthetic DNA or would they have the necessary
knowledge of relevant laboratory procedures? The failure to answer this question has led
the U.S. government and the scientific community to focus on three ways to deal with the
potential threat of DNA synthesis: (1) DNA sequence screening, (2) institutional review

\textsuperscript{116} David Relman, “Working Group on Synthetic Genomics: Progress Report” (report presented at a
meeting of the National Science Advisory Board for Biosecurity, July 13, 2006). An NSABB report
published later in 2006 identified two biosecurity concerns associated with synthetic genomics: (1) that it
could be used to synthesize and produce select agents and (2) that it could be used to produce agents that
resemble and have similar attributes to select agents without being clearly identifiable as a select agent
based on genetic sequence.

\textsuperscript{117} The National Research Council’s 2010 Sequence-Based Classification of Select Agents: A Brighter Line
report provides the results of the feasibility assessment the NSABB recommended. According to the report,
credibly predicting an organism’s pathogenicity, transmissibility, environmental stability, virulence, or
other phenotype using sequence information is not plausible with current technology, because these
predictions require a detailed understanding of multiple pathogen and host attributes.
of dual-use research, and (3) restrictions on publications. Although these tools address the threat potential of DNA synthesis, they also have major weaknesses, which are described in the following section.

**Screening Genetic Sequences**

One proposed method to prevent malefactors from acquiring select agent DNA is to require synthetic DNA providers to screen customers’ orders to identify select agent genes. Some DNA synthesis companies began screening orders in the early 2000s, prompted by the terrorist attacks of September 11, 2001, so by the time the NSABB working group’s recommendations were published in 2006, sequence screening was already widely accepted among gene synthesis companies in the United States.\(^\text{118}\)

However, the method for screening orders varied from company to company. Shortly after the NSABB report was published, a consortium of German companies formed the International Association of Synthetic Biology (IASB), which held an open workshop to create a code of conduct for screening customers and orders. The code, shown in appendix C, requires IASB member companies to check all incoming orders against the GenBank\(^\text{119}\) sequence database to identify sequences that might be hazardous. When a sequence is found to be hazardous,\(^\text{120}\) it is referred to a human screener who must determine whether the customers are legitimate.\(^\text{121}\) The IASB grants member firms a

\(^{118}\) J. Tucker, “Double-Edged DNA” (see chap. 1, n. 1). Blue Heron Biotechnology, for example, had been screening customers since its inception in 2001, and it began screening DNA orders later that year, after the anthrax letter attacks.

\(^{119}\) GenBank is a database containing all publicly available DNA sequences. It is maintained by the National Institutes of Health.

\(^{120}\) These hazardous sequences are often associated with pathogens on the select agent list or on the Australia Group’s common control lists of materials and technologies.

“seal of approval” that they can display publicly to show that they are reputable, safety-conscious suppliers, which might confer a competitive advantage.\textsuperscript{122}

In 2009, DNA2.0 and GeneArt proposed a new screening method that replaced human screeners with automated systems, a more efficient and cost-effective method. However, this method was not adopted because, due to limitations in automated screening software, it was considered less robust than human screening.\textsuperscript{123} In September 2009, a new coalition called the International Gene Synthesis Consortium (IGSC) was established, and it promptly released its Harmonized Screening Protocol, shown in appendix D. Like the IASB’s code of conduct, the IGSC’s protocol requires member companies to screen orders against various databases to determine whether the ordered materials are hazardous and to screen the legitimacy of customers as needed. IGSC members must also report to the FBI those customers who have ordered hazardous sequences and whose legitimacy cannot be confirmed. They must also retain records of all gene synthesis orders for eight years.\textsuperscript{124} IGSC members include U.S. industry leaders Gen9, DNA2.0, GenScript, Blue Heron Biotechnology (now a subsidiary of OriGene), Integrated DNA Technologies, SGI-DNA, and Thermo Fisher Scientific. These companies currently constitute 80 percent of the global gene synthesis industry.\textsuperscript{125}

\textsuperscript{122} Tucker, “Double-Edged DNA.”
\textsuperscript{123} Ibid.
The IASB and IGSC are very similar, requiring members to screen both sequences and customers, maintain records of all orders, and establish contacts with law enforcement agencies, whom they will contact if they encounter a suspicious order.\textsuperscript{126} The only substantial differences in the two screening standards are the way they were developed and the way the standards will likely look in the future. The IASB’s code was developed using input from any firm that wished to participate in the process, while the IGSC was developed in private by the firms who enjoyed the largest market share at that time. Because the IASB and IGSC were developed when automated screening technologies were not robust enough to replace human screeners, both standards currently use human screeners to check sequences that are potentially hazardous. However, IGSC members are likely to replace human screeners with automated screening technologies as soon as these systems have proven sufficiently capable.\textsuperscript{127}

Although two industry-developed screening standards already existed, the U.S. DHHS in November 2009 released a draft of its screening guidance: \textit{Screening Framework Guidance for Synthetic Double-Stranded DNA Provider}.\textsuperscript{128} The DHHS proposed guidelines do not require synthesis companies to screen their orders, but they recommend that companies implement a best-match approach that involves screening sequence orders against a publicly available database of all known nucleotide sequences, such as GenBank, to identify the top-matching genetic sequence. If the top match is a select agent sequence, then a human expert should perform follow-up screening to further

\textsuperscript{126} Tucker, “Double-Edged DNA.”
\textsuperscript{127} Ibid.
assess the order, such as conducting literature searches to verify customers’ research
focus or calling researchers to inquire further. The framework also recommends that
synthesis firms screen customers against several lists of proscribed entities, such as the
Department of Commerce’s Denied Persons List and the Department of State’s list of
persons engaged in proliferation activities. Firms who are shipping synthetic genes
outside of the U.S. are also required to comply with U.S. trade sanctions and export
controls. These requirements are similar to the IASB and IGSC protocols.

However, participants in a January 2010 scientific workshop hosted by the Center
for Science, Technology and Security Policy at the American Association for the
Advancement of Science agreed the best-match approach is weaker than the industry-
developed screening measures because it only identifies select agent sequences, not
sequences for nonselect agent pathogens.\footnote{129} The ability to identify nonselect agent
pathogens is also important for security, since terrorists might be interested in using a
pathogen that is not on the list—such as \textit{Legionella pneumophila}, the causative agent of
Legionnaire’s Disease, or \textit{Escherichia coli} O157:H7, the causative agent of
enterohemorrhagic \textit{E. coli}. These two organisms are not currently on the select agent list,
but they are included in the Australia Group list for export control.\footnote{130} Additionally,
malefactors might also wish to use emerging pathogens, which must first be assessed
before they are added to the select agent list, a time-consuming process. For example, the
coronavirus that was responsible for the severe acute respiratory syndrome (SARS)

\footnote{129} Tucker, “Double-Edged DNA.”
\footnote{130} \textit{E. coli} O157:H7 is a shiga-like toxin, which was previously included in the select agent list but is not in
the current October 22, 2015 version.
pandemic in 2003 was not added to the select agent list until 2012.\textsuperscript{131} This delay was partially due to concerns that restricting access to the SARS coronavirus would have hindered medically important research. In 2004, more than a dozen scientists asked the CDC to keep the SARS coronavirus off the select agent list for this reason.\textsuperscript{132}

Unlike the DHHS framework, which was finalized in 2010, the standards developed by the IASB and IGSC require screeners to investigate every gene shipped to customers against multiple control lists—not just the Select Agents and Toxins list—to identify hazardous sequences.\textsuperscript{133} This review process is more thorough than the DHHS framework because it allows screeners to identify a broader range of pathogen and toxin sequences; however, it is also more expensive and time-consuming. Indeed, a recent study revealed that approximately five percent of orders that IGSC members receive require investigation, and these investigations can take up to 90 minutes to resolve.\textsuperscript{134} Moreover, if a legitimate firm outside of the United States purchases a gene associated with an export-controlled organism from a U.S. synthesis firm, the synthesis firm must license the order. This process could take several months, which might cause customers to purchase genes from companies in countries that do not screen orders.

Despite the widespread acceptance of sequence screening among U.S. gene synthesis companies, the utility of screening to prevent illegitimate access is dependent upon numerous factors, such as the completeness of databases containing nucleotide


\textsuperscript{133} DHHS, “Screening Framework Guidance.”

\textsuperscript{134} Carter and Friedman, “DNA Synthesis and Biosecurity.”

62
sequences, as well as the vigilance and acumen of human screeners. For example, the nucleotide sequences for rapidly mutating ribonucleic acid (RNA) viruses and pathogens that cause emerging infectious diseases are not readily available in public databases because they are so new they have not been sequenced yet. This implies that in cases where sequences are not in existing databases, human screeners must attempt to discern the nature of the customer’s research. This can be a time-consuming task for synthesis firms, requiring much research and investigation to determine whether customers have a legitimate research need for the sequences.

In October 2015, the JCVI released a report describing how well DNA synthesis firms have implemented the DHHS’ screening guidance. The report concluded that while existing screening guidance is working “reasonably well,” the availability of benchtop gene synthesizers could decrease synthesis costs and decentralize synthetic gene production in the future, making it more difficult for providers to allocate resources toward complying with the DHHS guidance.\textsuperscript{135} The JCVI report offered two recommendations for ensuring screening guidance remains relevant in the future. First, it recommended that the U.S. government require federal grantees to purchase genes only from companies that comply with the DHHS guidance. This option would require DNA providers to obtain a certificate from the government showing that they comply with the DHHS guidance. Second, it recommended that the U.S. government develop databases and software tools to help companies screen orders for sequences of concern. Specifically, the report recommended that the government prioritize sequences that

\textsuperscript{135} Ibid., 16.
“moderately skilled” scientists could make pathogenic, such as those that are associated with organisms with small genomes.\textsuperscript{136} The purpose of the two recommendations is to increase the number of companies that comply with the DHHS guidance, as well as to make complying with the guidance less burdensome.

\textbf{Conducting Institutional Reviews of Synthetic Biology Experiments}

In addition to the screening methods that aim to prevent malefactors from obtaining genes from harmful pathogens and toxins, several policies exist to improve biosafety in laboratories that conduct dual-use research. In 2007, the NSABB described a proposed framework for providing oversight of federally funded and federally conducted dual-use life sciences research. The proposal urges the government to require researchers and institutions that receive federal funding for potentially dual-use life sciences research to (1) receive training on dual-use concerns, (2) periodically review their research to determine its dual-use potential, and (3) take measures to reduce and manage risks if the research has dual-use potential.\textsuperscript{137} The NSABB also urged the government to develop and implement an institutional review policy for government-funded and -conducted dual-use research, since no such policy existed.\textsuperscript{138} The NSABB repeated these recommendations in a 2010 report that addressed the biosecurity and dual-use concerns arising from synthetic biology.\textsuperscript{139} In the 2010 report, the NSABB recommended that synthetic biology

\textsuperscript{136} Ibid., 18.
\textsuperscript{138} Ibid.
\textsuperscript{139} NSABB, “Synthesis of Select Agents”
research be subject to institutional review and oversight and that this oversight should extend beyond the life sciences and academia to other fields, such as public health and agriculture. In addition, the NSABB recommended that the U.S. government work with the scientific community to ensure that current biosafety guidelines associated with synthetic genomics are appropriate. In response to these recommendations, the National Institutes of Health (NIH) updated its NIH Guidelines for Research Involving Recombinant DNA Molecules—hereafter called NIH Guidelines—in 2013.\textsuperscript{140}

The NIH Guidelines specify principles for constructing and handling recombinant nucleic acid molecules, such as RNA and DNA; synthetic nucleic acids, such as those that are produced chemically; and cells and organisms containing recombinant or synthetic nucleic acids.\textsuperscript{141} Before the NSABB published its recommendations, the NIH Guidelines had only addressed recombinant nucleic acids. Per the NIH Guidelines published in 2013, institutions that receive NIH funding for any research must establish an institutional biosafety committee to review research involving recombinant or synthetic nucleic acids.\textsuperscript{142} Institutions that receive NIH funding are also required to appoint a biological safety officer to oversee the management of biosafety risks, if the institution is engaged in large-scale research or production activities with recombinant or synthetic molecules. Exempt from NIH Guidelines are synthetic nucleic acids that (1) cannot replicate or generate nucleic acids that can replicate in a living cell (such as oligonucleotides, which are typically single-stranded and not functional); (2) do not


\textsuperscript{141} Ibid.

\textsuperscript{142} Ibid.
integrate into DNA; (3) do not produce a toxin that is lethal in vertebrates at a certain dose; and (4) are not contained in a biological system such as a cell, organism, or virus.\textsuperscript{143} This latter stipulation exempts chemical synthesis of DNA from the NIH Guidelines until that synthetic DNA is placed in a cell or organism. The NIH Guidelines apply to institutions that receive NIH funding for any nonexempt research involving recombinant or synthetic nucleic acids. Institutions that receive NIH funding for research with these molecules will need to comply with the 2013 NIH Guidelines if they are simultaneously performing research with these molecules for other customers—NIH or otherwise.\textsuperscript{144} Other federal agencies, such as the Departments of Agriculture and Energy, require that research they fund involving the use of recombinant or synthetic nucleic acids also comply with the guidelines.

To address the NSABB’s recommendations that the government develop an institutional review process for government-funded or -conducted dual-use research, the U.S. government released its \textit{Policy for Oversight of Life Sciences Dual Use Research of Concern} (DURC) in March 2012.\textsuperscript{145} The purpose of the policy is to provide a framework by which government agencies can mitigate the risk that any dual-use research they fund or conduct be misused. The policy articulates 15 high-consequence pathogens and toxins

\begin{footnotesize}
\begin{enumerate}
\item[Ibid.]
\item[Ibid.]
\item[Office of Science and Technology Policy, “United States Government Policy for Oversight of Life Sciences Dual Use Research of Concern,” United States Government 2012, http://osp.od.nih.gov/sites/default/files/resources/United_States_Government_Policy_for_Oversight_of_DURC_FINAL_version_032812_1.pdf. The policy defines DURC as life sciences research that, based on current understanding, can be reasonably anticipated to provide knowledge, information, products, or technologies that could be directly misapplied to pose a significant threat with broad potential consequences to public health and safety, agricultural crops and other plants, animals, the environment, materiel, or national security.]
\end{enumerate}
\end{footnotesize}
and seven categories of DURC. These categories include experiments that (1) enhance the harmful consequences of the 15 agents or toxins; (2) disrupt host immunity or the effectiveness of vaccines against them without any clinical or agricultural justification; (3) render the agents or toxins resistant to prophylaxes or other medical interventions or allow them to evade detection techniques; (4) improve the stability, transmissibility, or ability to disseminate the agent or toxin; (5) alter the host range or tropism to the agent or toxin; (6) enhance host susceptibility to the agent or toxin; or (7) generate eradicated agents or toxins. According to the policy, federal departments and agencies that conduct or fund these types of experiments must assess the risks and benefits of the projects—including how the knowledge and technologies that result from the projects increase risk—and develop and implement a risk-mitigation plan with appropriate risk-mitigation measures. These measures could include enhancing biosecurity and biosafety measures, determining how to communicate the research and its results responsibly, and regularly reviewing emerging research findings to identify additional DURC, among other measures. However, the policy only addresses DURC that is conducted or funded by the U.S. government, leaving commercial and private DURC unaddressed.

In 2014, the government released a report to strengthen institutional review of DURC: *Policy for Institutional Oversight of Life Sciences Dual Use Research of*  

---

146 These pathogens and toxins include the following: avian influenza virus (highly pathogenic), *Bacillus anthracis*, botulinum neurotoxin, *Burkholderia mallei*, *Burkholderia pseudomallei*, Ebola virus, foot-and-mouth disease virus, *Francisella tularensis*, Marburg virus, reconstructed 1918 influenza virus, rinderpest virus, toxin-producing strains of *Clostridium botulinum*, variola major virus, variola minor virus, and *Yersinia pestis*. All 15 are included in the current select agent list and/or the Australia Group’s List of Human and Animal Pathogens and Toxins for Export Control.
The policy applies to (1) U.S. government departments and agencies that fund or conduct life sciences research; (2) institutions within the United States which conduct or sponsor life sciences research that is funded by the U.S. government and which conduct or sponsor research that involves one or more of the 15 agents or toxins described in the March 2012 policy, even if that research is not funded by the U.S. government; and (3) institutions outside of the United States that are conducting research involving one or more of the 15 agents or toxins and that receive U.S. funding or sponsorship. The policy describes the roles and responsibilities of principal investigators, government-funded institutions, and the government’s funding agencies in identifying DURC and taking measures to reduce risks. It also articulates the process for institutional review of life sciences research. One of the responsibilities that institutions bear is to establish an institutional review entity who is responsible for reviewing research that is within the scope of the policy. Failure to comply with the policy could result in an institution losing its government funding. Institutions that conduct DURC involving the 15 agents or toxins identified in the policy but do not receive government funds are not subject to the policy’s oversight; however, the government strongly encourages these institutions to implement internal oversight procedures.

Office of Science and Technology Policy, “United States Government Policy for Institutional Oversight of Life Sciences Dual Use Research of Concern,” United States Government, 2013, http://www.phe.gov/s3/dualuse/documents/durc-policy.pdf. The policy notes that institutional oversight is a critical component to mitigating threats associated with DURC, because institutions are more familiar with the research that is being conducted in their facilities, and are better able to identify security risks and determine how best to mitigate those risks. The policy defines “institutions” as any government agency (Federal, State, tribal, or local), academic institution, corporation, company, partnership, society, association, firm, sole proprietorship, or other legal entity conducting research.
While the government and institutional reviews described in the 2013 NIH Guidelines and the guidance on DURC are useful tools for identifying and mitigating risks, they have limited applicability. The NIH Guidelines are not used by all government agencies that fund research using synthetic DNA. Moreover, neither the NIH Guidelines nor the government’s policies on DURC apply to privately funded research. Therefore, the policies would not likely prevent terrorists or other malefactors from misusing synthetic genomics or synthetic biology. In addition, the mere possession of an institutional review board (IRB) does not guarantee that a dual-use technology will not be misused because the IRB’s oversight might not be implemented in such a way as to prevent a technology from being misused. For example, an IRB might call for increased biosafety measures within its institution, such as establishing regular audits of a laboratory’s inventory, but laboratory personnel might not conduct the audits regularly enough to prevent theft or identify it in a timely manner. Similarly, an IRB might recommend restricting access to laboratories that are engaged in infectious diseases research, but that would not prevent a scientist working in that laboratory from accidentally or deliberately releasing one of the laboratory’s pathogens. There is also evidence that IRBs might lack the knowledge to identify DURC in the first place; indeed, many IRBs do not include people with extensive knowledge of biosafety or national security issues. In addition, IRBs typically focus on DURC materials that might be misused, not knowledge.

Censoring Scientific Literature
In addition to institutional reviews of synthetic biology experiments, several organizations have suggested that the scientific publications resulting from such experiments should be censored to exclude the materials and methods sections. The 2002 poliovirus synthesis prompted several academic journals—including Science and Nature—either to consider modifying or censoring articles they believed could compromise national security or to request that the authors self-censor.149 The NSABB in 2012 requested that an article not be published because it described how the highly pathogenic Asian avian influenza A (H5N1) virus could be transformed to enhance its transmissibility; however, the NSABB later rescinded the request.150

These efforts to censor scientific publications suggest that many believe malefactors could use a scientific article’s materials and methods section as a blueprint to create a synthetic organism. However, malefactors are unlikely to have the technical expertise necessary to replicate scientists’ methods using the information found in these sections. Although synthetic biologists have made significant progress in creating organisms using synthetic genes and genomes, they often encounter technical difficulties, many of which have long plagued genetic engineers.

Some of these difficulties, which I describe in greater detail in the following chapter, are associated with incomplete or vague language in laboratory protocols, including the materials and methods section in published literature. Journals tend to impose length restrictions on their materials and methods sections, which may result in

149 Harmon, “Journals to Consider U.S. Security” (see chap. 1, n. 10).
the omission of critical information. Additionally, some academic articles are deliberately vague so that researchers can obscure information that might give them a competitive advantage over other researchers.\textsuperscript{151} Also, some critical information is often missing from these articles because it cannot be easily translated into words. Current evidence shows that terrorist attempts to develop biological agents using published information have failed.\textsuperscript{152} Thus, while it is doubtful that censoring scientific publications can prevent malefactors from acquiring the knowledge necessary to use synthetic genomics or synthetic biology, such action can nonetheless stifle developments in important fields, such as medical diagnostics and pharmaceuticals, because scientists rely on the literature to stay current in their fields, optimize laboratory methods, and identify candidates for applied research. Therefore, any censoring policy that aims to prevent the misuse of these technologies will also have an impact on scientific innovation.

Many of the mitigation measures described above were put in place because of fears identified by threat assessments, which often characterize biotechnologies as easy to use so long as individuals have access to materials and explicit forms of knowledge\textsuperscript{153} such as published protocols. However, STS literature, as well as studies that examine knowledge management and innovation in the biotechnology industry, reveal that there are a number of contingencies—unexpected technical difficulties—associated with scientific work. The studies reveal that access to materials and explicit knowledge alone


\textsuperscript{153} Explicit knowledge is knowledge that is codified in the form of scientific protocols, instructions, and procedures.
are not sufficient to enable scientists to replicate synthetic biology experiments. Rather, success in these experiments relies heavily on tacit knowledge, which sociologist Harry Collins has defined as “knowledge or abilities that can be passed between scientists by personal contact but cannot be set out or passed on in formulae, diagrams, or verbal descriptions and instructions.”\textsuperscript{154} Because tacit knowledge is stored in individual scientists’ brains, it is difficult to transfer, often requiring direct interaction. The next chapter delves more deeply into STS research findings.

CHAPTER 3
STS LITERATURE: EVIDENCE OF EXPERIMENTAL CONTINGENCIES

STS literature, as well as research in other fields, describes contingencies that scientists often encounter when conducting various experiments in the biological sciences. Detailed interviews with scientists who conduct research provide valuable insight about the technical hurdles that terrorists or other malefactors could expect to encounter if they tried to use biotechnologies for nefarious purposes. This chapter will focus on the contingencies identified in current STS literature that pertain to three technologies that malefactors trying to produce a pathogen would have to use: PCR, DNA synthesis, and synthetic biology. PCR is a widely used technique that is crucial for the successful completion of many biological experiments, including synthetic biology. DNA synthesis—in which gene- or genome-length DNA is chemically produced—is a key enabler for synthetic biology and would be a required process. Synthetic biology, in which pathogens are created from a chemically derived genome, is the means by which malefactors would create pathogens from synthetic DNA.

STS literature shows that these three techniques are not as automated and standardized as is commonly believed and they often pose a number of unexpected problems that scientists need to solve as their experiments progress. The studies also show that scientists must have specific expertise that has been acquired over a long
period of time to solve the difficulties that arise. In fact, sometimes the problems cannot be solved; even expert scientists do not always possess the requisite knowledge, despite their years of experience.

This chapter also summarizes STS literature that describes how individuals and organizations acquire and transfer knowledge. These studies provide crucial insight into how terrorists or other malefactors might acquire the knowledge that is necessary to overcome experimental contingencies. Research shows that explicit forms of knowledge—such as the materials and methods sections of scientific literature—do not convey all of the information necessary to successfully complete an experiment because they omit the authors’ tacit knowledge. Acquiring this tacit knowledge requires hands-on training and practice. Thus, published protocols are insufficient blueprints for replicating synthetic biology processes.

**Contingencies in Life Science Research**

STS literature provides a number of examples of how scientists perform various techniques, as well as the difficulties they encounter. The following sections describe the contingencies that malefactors would encounter when performing PCR, synthesizing DNA, and conducting synthetic biology. Malefactors would need to use all three of these techniques to create a pathogen using synthetic biology.

**Contingencies with PCR**

Most molecular biology and genetic experiments involve the study of isolated pieces of DNA, so scientists must obtain many copies of the DNA pieces they intend to study. PCR, which is a technology used to copy small pieces of DNA millions of times, is
essential to carrying out these experiments. Indeed, NIH estimates that studying isolated pieces of DNA would be “nearly impossible” without PCR amplification.¹ Developed in 1983 by Kary Mullis, who went on to win a Nobel Prize in Chemistry for the invention, the technique has been used widely for more than 25 years, thanks in part to the availability of automated PCR kits that streamline the technique by providing scientists with ready-made reagents and a protocol that is purportedly easy to read.² These kits are frequently used in the life sciences because they enable scientists to do automatically what would otherwise require special skills; in theory, a scientist should be able to conduct PCR without experiencing any problems.

However, STS researchers Kathleen Jordan and Michael Lynch conducted a study in 1998 to examine how different organizations used PCR; the research revealed that scientists often experienced difficulty performing PCR, even though the technique had been widely used for a decade and was believed to be essentially automated. These findings suggest that PCR is far more complicated than is commonly believed. In the study, Jordan and Lynch interviewed technicians, staff scientists, laboratory administrators, and others who used PCR; the interviewees admitted that they often experienced great difficulty successfully using the technique, despite the availability of

² For an example of a typical PCR kit, see the Qiagen Multiplex PCR Kit: https://www.qiagen.com/us/shop/assay-technologies/end-point-PCR-and-RT-PCR-reagents/qiagen-multiplex-PCR-kit
step-by-step protocols as well as commercial kits intended to streamline the process.\(^3\) This difficulty—which some of Jordan and Lynch’s interviewees called “PCR hell”—prompted many of the scientists to conduct lengthy trial-and-error, modify the original protocol, make their own reagents, or solicit advice from other scientists.\(^4\) Consequently, Jordan and Lynch received remarkably different accounts from interviewees about how they conducted PCR—even from staff members within the same organization. Due to these factors, Lynch concluded that PCR is procedurally flexible; that is, its protocols vary depending on the sequence that is to be amplified, the equipment and reagents at hand, and other considerations.\(^5\)

Furthermore, Jordan and Lynch’s work also suggested that even though protocols exist for certain techniques, they do not describe precisely what scientists do. Instead, the protocols assume the scientists already possess certain competencies, such as prior experience using a technique or the ability to comprehend what is meant by vague laboratory language, such as “gently” pipetting a sample or adding “approximately” a certain amount of reagent.\(^6\) Moreover, Jordan and Lynch found that a PCR protocol that is successful in one circumstance may not be successful in another, prompting Lynch to conclude that “no single formulation of the PCR protocol gives a precise description of the range of techniques performed under its name, nor does it exhaustively describe the

---

3 Jordan and Lynch, “The Dissemination, Standardization, and Routinization” (see chap. 1, n. 11). Jordan and Lynch define protocols as written instructions that specify the ingredients, equipment, and sequence of steps for making technical preparations.

4 Ibid., 786.


6 Ibid., 206.
actions involved in a singular attempt to reproduce the technique.” 7 Jordan and Lynch’s research illustrates that scientific techniques involve many kinds of contingencies, as well as various processes and skills that must be continually adapted for new and different experiments. Additionally, their research shows that PCR is only theoretically automated. Because scientists must often modify protocols to each particular experiment, PCR cannot successfully be reduced to a standard protocol and reagent kit that can be used for any experiment.

**Contingencies in Oligonucleotide Synthesis**

In addition to the contingencies associated with PCR, literature from STS and other fields suggest that there are also contingencies associated with DNA synthesis. Oligonucleotide synthesis involves producing short fragments of DNA that are often single-stranded and not functional; the method was first used in academic laboratories in the 1950s, and the technique became commercialized and automated in the 1980s. 8 While there is little information about contingencies associated with oligonucleotide synthesis, a paper authored in 2005 by researcher Yogesh S. Sanghvi from Rasayan Inc.—commissioned for inclusion in *Synthetic Genomics: Risks and Benefits for Science and Society*—provides some detail about contingencies individuals might encounter if they synthesized oligonucleotides manually, without an automated synthesizer. 9 Specifically,

---

7 Ibid., 208.
9 Sanghvi, “Assembly of Synthetic DNA,” 28 (see chap. 2, n. 4). Sanghvi notes that DNA synthesis without an automated synthesizer is possible but highly inconvenient.
Sanghvi highlights contingencies associated with isolating nucleosides from natural sources; chemically synthesizing nucleosides; and storing and handling P(III) amidites, which are compounds that are derived from nucleosides.

Nucleotides, which are the building blocks of DNA, are the key reagents used in DNA synthesis. They can be ordered at low cost from several suppliers, or they can be made by adding phosphate groups to nucleosides. Nucleosides can be purchased for low cost from several suppliers, isolated from natural sources, or produced chemically. According to Sanghvi, the tasks of isolating nucleosides from natural sources and producing them chemically are labor-intensive and require significant expertise in chemistry—especially the chemical synthesis process—as well as specialized equipment. As such, an individual or group that is unable to purchase nucleosides will experience difficulty producing them from scratch.

Sanghvi also reports that individuals or groups who possess an automated DNA synthesizer but lack access to nucleotide amidites will have problems producing, handling, and storing these materials. The ability to successfully synthesize DNA using automated synthesizers depends on the quality of amidites. The production of high-quality amidites is “a skilled task,” according to Sanghvi, which requires large quantities of anhydrous solvents that must be kept in airtight containers. Sanghvi also asserts that isolating, storing, and handling P(III) amidites is “an art that is not easily acquired even

---

10 Sanghvi, “Assembly of Synthetic DNA.” Nucleosides are compounds that are commonly found in DNA or RNA. They consist of a nitrogenous base that is attached to a sugar but does not have a phosphate group. Nucleotides, by contrast, consist of a nitrogenous base and a sugar and have one to three phosphate groups.
11 Ibid.
12 Ibid., 22. Anhydrous solvents are water-free solvents.
by an experienced chemist.” However, he contends that a chemistry expert could produce large quantities of amidites in about six months, provided they have access to the necessary tools and chemicals, as well as training.

In addition to Sanghvi’s report, a 2014 article by synthetic biologists Sriram Kosuri and George Church describes the benefits and limitations of commonly used methods for gene synthesis. According to the researchers, although improvements in raw materials, automation, processing, and purification have enabled scientists to routinely synthesize up to 100 nucleotides (nt), there are limits to the length of the oligonucleotides that can be synthesized with the current methods. In addition, error rates are high, even for oligonucleotides that have been successfully synthesized.

Together these studies indicate that synthesizing oligonucleotides without an automated DNA synthesizer would be a difficult and time-consuming task. While an automated synthesizer would improve the process, it requires access to certain raw materials—nucleosides or amidites—without which a terrorist or any malefactor would encounter significant difficulties in synthesizing DNA. Moreover, any malefactor who succeeds in overcoming these hurdles would only be able to synthesize oligonucleotides of a certain length, and some of those would contain errors. Without appropriate expertise, malefactors will be unlikely to resolve the issues they face.

**Contingencies in Synthetic Genomics**

A review of STS literature shows that scientists also experience difficulties when synthesizing genes and genomes. Whereas oligonucleotide synthesis primarily relies on

---

13 Ibid., 22–23.
14 Kosuri and Church, “Large-Scale de Novo DNA Synthesis.”
chemistry, gene and genome synthesis—otherwise known as synthetic genomics—relies on molecular biology techniques and therefore requires specific expertise that is not required for oligonucleotide synthesis. Additionally, synthetic genomics require a more complex manufacturing operation compared to oligonucleotide synthesis because the sequences are longer, the behavior of the DNA is unpredictable, and there are high failure rates in the process. Although little literature exists that describes the difficulties associated with gene and genome synthesis, several disparate studies provide insight about limitations in the gene synthesis processes, the feasibility of constructing viral genomes using synthetic DNA, and contingencies encountered by JCVI researchers during the assembly of the approximately 583,000 bp *Mycoplasma genitalium* genome in 2008.

The previously mentioned 2014 article by Kosuri and Church on de novo DNA synthesis methods describes the benefits and limitations of commonly used methods for gene synthesis, including those used on a commercial scale. According to the researchers, polymerase cycling assembly (PCA)-based techniques, which are commonly used for gene synthesis, have high error rates. Moreover, sequences with high guanine-cytosine (GC) content and secondary structures can inhibit assembly. These results indicate that, despite the high level of automation in gene synthesis, the efficiency and success of the process is highly dependent on the nature of the DNA sequence being synthesized.

---

15 Vogel, *Phantom Menace*, 100 (see chap. 1, n. 11).
16 Kosuri and Church, “Large-Scale de Novo DNA Synthesis”; and Ibid., 100–101.
17 Kosuri and Church, “Large-Scale de Novo DNA Synthesis.”
In 2006, researcher Marc Collett explored the possibilities of constructing select agent viruses from synthetic DNA. Collett noted that, while it should be technically feasible for malefactors to create a viral threat agent from synthetic DNA, they would likely face technical challenges. For example, terrorists might encounter problems when trying to produce viral threat agents using reverse genetics techniques, because those techniques do not yet exist for some of the viruses. Consequently, malefactors would need to find alternative methods or develop new methods.

In 2013, science and technology researcher Kathleen Vogel described the difficulties that JCVI researchers encountered when synthesizing the approximately 583,000 bp *M. genitalium* genome in 2008. At the time, *M. genitalium* was the largest piece of synthetic DNA and the first bacterial genome ever assembled synthetically. The JCVI team used the techniques that were developed for assembling the phiX bacteriophage genome as a starting point for assembling the *M. genitalium*. The JCVI team asked three gene synthesis companies—Blue Heron Biotechnology, DNA2.0, and GeneArt—to synthesize 101 cassettes that were between 5,000 and 7,000 bp in length.

---

19 Ibid.
20 Reverse genetics refers to the development of molecular clones of infectious genomes. Reverse genetics systems are methods for determining the function of an organism’s genes and the phenotypes that are produced by mutations in those genes when all of the genes in an organism are known. Reverse genetics systems allow researchers to manipulate microorganism genomes and directly evaluate the effects of these manipulations.
22 Vogel, *Phantom Menace*.
Next, JCVI researchers assembled the cassettes into ever-larger fragments in five stages.\textsuperscript{24}

Vogel’s research indicates that the experiment took several years to complete and was a tedious process that involved building the genome a quarter at a time.\textsuperscript{25} Indeed, a Discovery Channel documentary about JCVI’s research indicated that, after more than a year of effort, the scientist responsible for assembling the \textit{M. genitalium} genome could only assemble half of the bacterium’s chromosome.\textsuperscript{26} Speaking in this documentary, Venter indicated that much of this difficulty was due to the novelty of working with such large fragments of DNA, which are brittle. According to Venter, there were no mechanisms for handling and manipulating chromosomes outside of cells, so the JCVI team had to develop those mechanisms as they went via trial and error.

A JCVI press release from January 24, 2008 explained that there 17 researchers on the team and that they received additional help from three gene synthesis firms.\textsuperscript{27} The team initially attempted to build the genome in \textit{Escherichia coli} but experienced difficulty, so they built it in \textit{Saccharomyces cerevisiae}—a yeast—instead. Although the JCVI team eventually succeeded in building the synthetic \textit{M. genitalium} genome, they noted that their method was prone to error. The sources of these errors, however, were difficult to pinpoint. The researchers speculated that they could have inadvertently

\textsuperscript{25} Vogel, \textit{Phantom Menace}, 99.
\textsuperscript{26} Science Channel, “Creating Synthetic Life.”
supplied Blue Heron, DNA2.0, or GeneArt with incorrect sequences, or the companies might have inadvertently delivered DNA fragments that contained errors. It is also possible that the errors occurred when JCVI researchers attempted to assemble the genome in *E. coli* or the yeast.\(^{28}\)

The researchers conducted quality control experiments throughout the experiment to ensure they could proceed to the next step, a time-consuming process. When asked about the experiment during a 2009 interview with *Nature Biotechnology*, Venter stated that researchers “have to solve each of these riddles [experimental contingencies] one at a time. That’s why it’s so slow.” He further explained that, in the case of *M. genitalium*, his team “just lucked out with one system. . . . It just as easily could have led us down a blind alley for a long time.”\(^{29}\) In sum, this case study indicates that synthesizing large genes and genomes is difficult and requires specialized skills, some of which are acquired as the experiment is being conducted.

While the literature on synthetic genomics is limited, there is ample data about contingencies associated with two synthetic biology experiments: the de novo synthesis of poliovirus in 2002 and the synthesis of the phiX bacteriophage in 2003. This literature indicates that a significant amount of skill was required to perform these experiments and, most importantly, the scientists who performed the poliovirus experiment were rarely able to replicate the experiment years after the protocol had been published.


\(^{29}\) Marshall, “Sorcerer of Synthetic Genomes,” 1123 (see chap. 1, n. 12).
Contingencies in Synthesizing the Poliovirus: An STS Case Study

In 2002, Eckard Wimmer and colleagues at the State University in New York at Stonybrook published an article describing the de novo synthesis of the 7,500 nt poliovirus, demonstrating that viruses could be synthesized in the absence of previously infected cells by using the genetic sequence of the pathogen of interest.\(^\text{30}\) This new revelation prompted many analysts to speculate whether malefactors or terrorist organizations could adapt the protocol to synthesize viruses such as Ebola or smallpox without obtaining stocks of the viruses through commercial or other means.\(^\text{31}\) Shortly after the online publication of Wimmer’s study, some scientists and policymakers claimed the experiment itself, as well as the resulting publication, was irresponsible from a national security standpoint.\(^\text{32}\)

Although some scientists and policymakers believed a malefactor could use the experimental protocol to produce a synthetic pathogen, research by Vogel indicates that the experiment was fraught with contingency and was difficult even for Wimmer and his colleagues to replicate in the laboratory years after the study was published.\(^\text{33}\) According to Vogel (2008), the most difficult part of the experiment was the production and maintenance of the HeLa cell-free cytoplasmic extracts in which viral synthesis would occur.\(^\text{34}\) Producing these extracts involved growing HeLa cells\(^\text{35}\) in cell culture, manually

\(^{30}\) Cello, Paul, and Wimmer, “Synthesis of Poliovirus cDNA” (see chap. 1, n. 5).
\(^{33}\) Vogel, Phantom Menace, 76.
breaking them up with a glass Dounce homogenizer after the cells were removed from
culture, removing the HeLa cells’ nuclei and other components, dialyzing the resulting
HeLa cell-free cytoplasmic extract, and eventually storing the cell-free extract until it was
needed for the experiment.

Vogel’s interviews with Wimmer and his colleagues revealed that although the
procedures for developing and maintaining these extracts had been available for more
than 20 years and various papers provided detailed instructions of how to achieve viral
synthesis, the act of producing and maintaining the extracts was extremely difficult to
reproduce. According to Vogel, even though the researchers used the same protocols
and reagents each time they attempted to make the extracts, their ability to make the
extracts successfully varied from day to day. Notably, the researchers did not know why
they could make the extract successfully one day but not the next, and they estimated that
only about one in four batches were successful. The reasons for this are unknown, despite
the fact that Wimmer’s coauthors, Dr. Aniko Paul and colleagues attempted to identify
the source of the variation and always used the same procedure.

Vogel’s interviewees proposed a variety of possible contingencies that may have
affected their ability to reproduce the experiment, including variations in reagents and
materials provided by suppliers. One researcher believed that the contingency might be
due to variations in the fetal bovine serum that was used to grow HeLa cells for
poliovirus synthesis. Specifically, the researcher suggested that variations in different

---

35 HeLa cells are immortal cells often used in scientific research.
36 Vogel, Phantom Menace, 77.
37 Ibid.
38 Ibid.

85
companies’ serum might influence the amount of serum researchers would need to use. Another researcher suggested that the quality of the bovine serum might vary with the seasons. Researchers also surmised that variations in nucleases provided by suppliers might have affected the experiment’s reproducibility.

The researchers further speculated that water quality might affect the success of the experiment. According to Vogel’s interviews, a Belgian graduate student who made successful extracts in the Wimmer laboratory was never successful making the same extracts when he returned to Belgium, even though he followed the Wimmer laboratory protocol. This led to speculation that the water in Belgium might have a different pH or mineral content, which could potentially affect the reagents used in the experiment. This example indicates that tacit knowledge is local; when taken out of context, it is not sufficient to master a scientific experiment.

Another problem in preparing cell-free extracts was the use of the Dounce homogenizer, which is a glass tube with a pestle that fits inside. To break the HeLa cells open and release the cytoplasmic extract, HeLa cells are placed into the glass tube and the pestle is moved up and down. Wimmer and colleagues described this process as an art in which the cells must be broken up just enough but not too much—something that is difficult to convey in a protocol. According to Vogel, researchers who have successfully used the Dounce homogenizer report that using it should feel the right way and look the right way—there are a number of physical, visual, and sensory cues that they pay

---

39 Vogel, Phantom Menace, 79.
40 Ben Ouagrham-Gormley, Barriers to Bioweapons (see chap. 1, n. 11); and Ibid., 79.
attention to as they use the homogenizer. Since there are variations in Dounces, the exact number of strokes it takes to break a given number of cells needs to be empirically determined in each lab using trial and error. Researchers interviewed by Vogel stated that they always use the same Dounce to reduce contingencies associated with using the instrument.

Vogel’s interviews reveal that there is a great deal of contingency associated with the poliovirus experiment, some of which cannot be overcome in a predictable way because there are a variety of social (skills, disciplines) and environmental (water, pH, nucleases) factors that can influence the environment. Even after having conducted detailed control experiments, the researchers cannot say for certain which factors are the most important for experimental success. In a July 2006 article in *EMBO Reports*, Wimmer summarized the reproducibility problem in biology:

> The outcome of an experiment in fundamental research biology . . . cannot be predicted. Biological systems are too complex to be able to account for all variations, no matter how carefully the experiment is designed. . . . Intuition and creativity play as great a role in science as they do in the arts, but are no guarantee of success. Thus, even the best scientists suffer frequent humiliation when their experiments do not yield the expected results.

---

41 Vogel, *Phantom Menace.*
42 Ibid.
Contingencies in Synthesizing the phiX Bacteriophage: An STS Case Study

More than one year after the results of the poliovirus synthesis were published, the JCVI published an alternative, quicker method for synthesizing a larger genome from a different organism, the phiX bacteriophage. The phiX bacteriophage infects *E. coli*, and its genome is approximately 5,300 nt in length. The JCVI study demonstrated the possibility of quickly synthesizing genomes, and it appeared to support the claim that biotechnological knowledge was proliferating at a rapid pace and becoming less difficult to exploit, thus reinforcing the prevailing claim that malefactors could use the technology to create select agents or other pathogens. However, research by Vogel reveals that, like the poliovirus experiment, the synthesis of phiX bacteriophage was fraught with contingencies that would likely stymie a nonexpert who tried to adopt the technology.

According to Vogel, although the phiX bacteriophage synthesis was carried out within a two-week period, work on the project began nearly seven years earlier, and one of the JCVI researchers had more than 40 years of experience working with the bacteriophage. This researcher made initial attempts to synthesize the phage in 1996 but was unsuccessful. Vogel’s research reveals that when JCVI researchers initially attempted to synthesize the phage in 1996, they were unable to obtain any plaques because the phage contained mutations and was not viable. JCVI researchers attribute this to errors in oligonucleotide synthesis, which was carried out by commercial firms.

---

44 Smith et al., “Generating a Synthetic Genome” (see chap. 2, n. 41).
45 Vogel, *Phantom Menace*.
46 Ibid., 92.
47 A plaque is a visible formation within a cell culture that indicates a virus, such as a bacteriophage, is replicating.
48 Ibid., 93.
researchers estimated that only about half of the molecules produced in automated DNA synthesizers—such as those used in their phiX bacteriophage study—would have the correct length and that the incorrect molecules would interfere with the assembly of the correct oligonucleotides or cause mutations in the assembled DNA.\textsuperscript{49} Synthesis error rates are so common that Craig Venter, founder of the JCVI, in an interview with Vogel, said, “That is the problem with the whole field [of synthetic genomics]. If the starting material was 100 percent accurate . . . you’d just put them together and everything would be simple.”\textsuperscript{50} In an interview in \textit{Nature Biotechnology} in 2009, Venter revealed that correcting synthesis errors took the JCVI team years to resolve.\textsuperscript{51} Venter went on to say that the 2003 paper describing the phiX synthesis contained “some nice elegant methods for doing repair in real time off of a correct strand” of DNA, but the researchers “still had to select clones and sequence them to ensure the correct order of bases.”\textsuperscript{52} Said Venter, the phiX synthesis and subsequent \textit{M. genitalium} synthesis took several years and required the team to develop a new methodology and tools “at every stage.”\textsuperscript{53}

Following the failures in 1996, JCVI researchers set aside the experiment and did not revisit it until 2002. When the team revisited the research, they added two scientists to the team, one of whom was a Nobel Laureate and both of whom the principal investigator described as experts in handling DNA. When these two scientists tried to replicate the original research, they were unable to assemble a full-length piece of DNA.

\textsuperscript{49} Smith et al., “Generating a Synthetic Genome.”
\textsuperscript{50} Vogel, \textit{Phantom Menace}, 94.
\textsuperscript{51} Marshall, “Sorcerer of Synthetic Genomics.”
\textsuperscript{52} Ibid., 1122.
\textsuperscript{53} Ibid., 1122.
Thus, the team had to develop a new process to do so.\textsuperscript{54} The new process involved purifying oligonucleotides by length, a time-consuming and laborious process.

Some national security analysts believe the speed with which JCVI researchers were able to synthesize the bacteriophage is proof of the rapid pace of biotechnological advances.\textsuperscript{55} However, Vogel’s interviews indicate that the speed with which the experiment was conducted depended on researchers’ expertise and high skill sets.\textsuperscript{56} Vogel’s research also revealed that there was contingency associated with using plaque assays, which allowed researchers to rapidly and visually identify fully infectious diseases. To perform the assay, researchers injected synthesized phage into agar plates of \textit{E. coli} and waited for large plaques to form on the plates. Large plaques typically signify accurate, infectious phage; however, the phiX experiment produced only one plaque-forming genome per 10,000 colonies.\textsuperscript{57} Terrorists who wish to replicate this study would also have a difficult time using plaque assays. Indeed, the inability to use a plaque assay would greatly increase the amount of time, complexity, and effort required to synthesize larger genomes.\textsuperscript{58}

In addition to Vogel’s case studies, the scientific literature confirms that although synthetic biologists have made significant progress in creating organisms containing synthetic genes and genomes, they are still encountering difficulties, including those that have long plagued genetic engineers. First, synthetic biologists have not been able to

\textsuperscript{54} Vogel, \textit{Phantom Menace}, 94.
\textsuperscript{55} Chyba, “Biotechnology and the Challenge” (see chap. 2, n. 45).
\textsuperscript{56} Vogel, “Framing Biosecurity.”
\textsuperscript{57} Vogel, \textit{Phantom Menace}, 97.
\textsuperscript{58} Ibid., 97.
reliably engineer biological systems with predefined properties.\textsuperscript{59} This limitation is due to incomplete knowledge of biological systems, as well as the fact that biological systems have a large number of functional components that interact in unexpected ways, exhibiting unpredictable behavior. Much work is still needed to enable researchers to predict biological characteristics from nucleic acid or protein sequences or structures.\textsuperscript{60} Second, natural biological systems are not optimized by evolution for engineering; instead, they are optimized to perform within their natural contexts. As such, engineered systems may not function when connected to each other in an artificial context.\textsuperscript{61} Third, synthetic biologists may encounter pleiotropic effects—undesirable characteristics that appear in an engineered organism along with desired characteristics—when attempting to create organisms.\textsuperscript{62} Additionally, interviews with former Soviet and U.S. bioweaponeers have revealed that manipulating some pathogen strains could compromise their effectiveness as weapons.\textsuperscript{63} All of these limitations could greatly influence malefactors’ ability to create effective pathogens using synthetic DNA, and they must be considered when developing policies meant to prevent the misapplication of gene synthesis and synthetic biology.

The experimental contingencies described in this section indicate that (1) numerous contingencies—which can result from conducting novel tasks, using different

\textsuperscript{59} Endy, “Foundations for Engineering Biology” (see chap. 1, n. 12).
\textsuperscript{60} National Science Advisory Board for Biodefense, \textit{Sequence-Based Classification of Select Agents: A Brighter Line} (Washington DC: National Academies Press, 2010).
\textsuperscript{63} Ben Ouagrham-Gormley, \textit{Barriers to Bioweapons}, 6.
laboratory instrumentation, or altering various environmental factors—arise when conducting scientific tasks; (2) researchers may not be able to discern the source(s) of contingency; (3) published protocols and other forms of explicit knowledge play a limited role in the success and replication of experiments; and (4) the success of these experiments relies heavily on researchers’ tacit knowledge—both knowledge that was previously obtained, as well as knowledge that is acquired during the trial-and-error process of experimental science. The importance of tacit knowledge indicates that malefactors would have difficulties replicating the protocols in published literature.

The following section describes the roles that explicit and tacit knowledge play in experimental work and summarizes STS literature to provide insight how individuals acquire and transfer tacit knowledge. The section also highlights factors that might facilitate or hinder the acquisition of tacit knowledge to envision how terrorists might acquire the necessary knowledge to create harmful pathogens using synthetic genomics or synthetic biology.

**The Role of Tacit Knowledge in Experimental Work**

STS literature shows that scientists can encounter significant complications when using biotechnologies, such as difficulty interpreting laboratory protocols, differences in equipment, and limitations associated with working with unpredictable living organisms. These nuances are missing from mainstream threat assessments, which assume that scientists can replicate previous experiments so long as they have access to materials and protocols. In contrast, STS literature demonstrates that, far from being “simple recipes,”

---

64 Carlson, “Pace and Proliferation,” 209 (see chap. 2, n. 1).
laboratory protocols, standard operating procedures, and other forms of explicit knowledge are insufficient for allowing individuals to replicate previous experiments. This is because these explicit forms of knowledge provide only knowledge that can be codified, and they omit knowledge that cannot be codified, such as tacit knowledge. Indeed, some of these explicit knowledge sources are purposely incomplete. For example, academic journals typically require authors to submit brief descriptions of their materials and methods, and this need for brevity inevitably results in the omission of critical information. In addition, protocols also use ambiguous language that can cause technical difficulties in the laboratory or make it difficult for scientists to use them to resolve problems.

Thus, STS literature reveals that explicit knowledge sources are procedurally flexible; that is, they change depending on the context in which scientists employ them. As such, scientists must customize explicit knowledge so that it can be applied to their experiments and settings. Kathleen Vogel’s case study of the 2002 poliovirus experiment provides an excellent example of the need to adapt explicit knowledge to the context of the recipient site: a Belgian graduate who had made successful HeLa cell-free cytoplasmic extracts in the Wimmer laboratory was not able to make them when he returned to Belgium, even though he had followed the Wimmer laboratory protocol to the letter.

---

65 Ben Ouagrham-Gormley, *Barriers to Bioweapons*, 27.
66 Jordan and Lynch, “The Dissemination, Standardization, and Routinization.”
67 Ibid.
68 Ben Ouagrham-Gormley, *Barriers to Bioweapons*, 27.
STS literature also demonstrates the crucial role that tacit knowledge plays in the laboratory sciences. Indeed, researchers rely heavily on tacit knowledge—either their own or that of their colleagues—to resolve technical difficulties. The STS literature indicates that much of this tacit knowledge is acquired through trial and error and repeating tasks frequently, or by learning in a master-apprenticeship relationship.\(^70\) However, it is unclear how easily tacit knowledge can be acquired or transferred. To provide insight about how malefactors might acquire the tacit knowledge necessary to create harmful pathogens using synthetic genomics or synthetic biology, the following section summarizes problems with transferring tacit knowledge and describes the most successful methods for transferring it among individuals.

**Problems Transferring Tacit Knowledge**

As mentioned previously, tacit knowledge is the unarticulated knowledge acquired through hands-on experience, which involves either learning by directly performing a task or by observing and collaborating with colleagues.\(^71\) STS literature has shown that scientists rely heavily on tacit knowledge to resolve difficulties. Yet, tacit knowledge is difficult to transfer because much of it can never be codified and because it is local in nature; in other words, knowledge that works in one context or setting may not work the same way in a different context.

Tacit knowledge is difficult to codify because some information is difficult to articulate, such as a description of the motor skills and series of movements that

---

\(^70\) Collins, “Tacit Knowledge, Trust, and the Q of Sapphire” (see chap. 2, n. 146).

\(^71\) Ben Ouagrham-Gormley, *Barriers to Bioweapons*, 27.
constitute gentle mixing. Additionally, scientists who codify their knowledge might unwittingly omit important information. In other words, they might possess what sociologist of science Harry Collins refers to as concealed knowledge. Scientist may not be aware of all of the knowledge they possess that is crucial for an experiment’s success (unrecognized knowledge), and they might not understand how they are able to conduct certain experiments or tasks successfully (uncognized and uncognizable knowledge).

As mentioned previously, tacit knowledge is also local in nature and must be adapted to be used in new contexts. A 2006 case study of the Soviet Union’s biological weapons facility in Stepnogorsk illustrates researchers’ need to adapt tacit knowledge to their local context. The Soviet Ministry of Defense gave the facility two tasks: improve the potency of Anthrax 836 and improve the organization’s ability to manufacture the weapon on a large scale. The Ministry of Defense had already developed a wool-based 836 formulation that embedded Bacillus anthracis spores in shredded wool. However, the wool material was heavy and would settle rapidly after dissemination, making the wool-based formulation unsuitable for military use.

Consequently, the Ministry of Defense tasked the Stepnogorsk facility to develop a new protocol for Anthrax 836, which used an encapsulation technology based on polyvinyl chloride, which is lighter than wool and is heat-protective; this encapsulation technology would increase the viability of the bioweapon following dissemination to

---

72 Collins, “Tacit Knowledge, Trust, and the Q of Sapphire.”
73 Ibid.
74 Ben Ouagrham-Gormley, Barriers to Bioweapons; and Vogel, Phantom Menace, 79.
75 Vogel, “Bioweapons Proliferation” (see chap. 1, n. 11).
adversaries. The Ministry of Defense gave Stepnogorsk’s scientists technical papers and plans pertaining to Anthrax 836 prior to the start of their work. These papers included a report summarizing how the wool-based strain was developed, information describing equipment and biosafety conditions for the production of the previous strain, instructions for filling and assembling bombs and warheads, and contact information for raw materials suppliers. However, despite being given these resources, the scientists had problems applying the information to their facility because it was specific to growing Anthrax 836 in Ministry of Defense equipment. As a result, the staff at Stepnogorsk could only use the Ministry of Defense documents as guidelines, and they had to develop new materials and methods that were specific to their facility. Many of the problems they encountered related to scaling up the Anthrax 836 strain from laboratory cultivation to large-scale production, as well as ensuring the equipment at the Stepnogorsk facility was capable of large-scale production, a process that involved testing and repairing thousands of pieces of equipment. To deal with these scientific and engineering problems, more than 200 new employees were sent to the facility within a year of it starting work, many of whom were scientists who had expertise in various aspects of weaponizing anthrax bacteria. According to Vogel, these experienced employees likely helped the original Stepnogorsk scientists adapt the Ministry of Defense protocols for use in their facility and provided them with valuable bioweapons training. However, even with this infusion of new employees, it took the experts an additional three years of work before they could
produce large quantities of the new anthrax strain, illustrating the difficulties of adapting tacit knowledge to new contexts.\textsuperscript{76}

Although tacit knowledge is difficult to transfer, research shows that it can be acquired through practice—either through trial and error or by directly observing and replicating the routines of other scientists.\textsuperscript{77} These methods are described in the following section and show how terrorists might acquire the expertise needed to create harmful pathogens using synthetic genomics or synthetic biology.

**Methods for Acquiring and Transferring Tacit Knowledge**

Security researcher Sonia Ben Ouagrham-Gormley provides a good overview of methods for transferring and acquiring knowledge.\textsuperscript{78} In this section, I discuss Ben Ouagrham-Gormley’s findings and complement her work with information from other studies.

The previous section shows that scientists acquire much of their tacit knowledge through trial and error. Research by Ben Ouagrham-Gormley indicates that the lessons learned from solving problems through trial and error are more likely to be remembered when they are applied in future endeavors, like performing tasks or transferring knowledge.\textsuperscript{79} A 2009 interview with Craig Venter revealed that the trial and error required to overcome contingencies in the *M. genitalium* and *M. mycoides* experiments


\textsuperscript{77} Ben Ouagrham-Gormley, *Barriers to Bioweapons*.

\textsuperscript{78} Ibid.

resulted in a better understanding of the fundamental scientific principles behind genome isolation and transplantation. According to Venter, he and his colleagues “learned so much that’s really critical for the next stages [just] by struggling to get through it.” This suggests that terrorists could acquire the expertise needed to create harmful pathogens using synthetic genomics or synthetic biology by meticulously working on experiments over time, testing slight variations by applying information that has been gleaned from the literature and more experienced scientists.

Another way that terrorists might acquire critical expertise is by directly observing and practicing the routines of other, more experienced scientists through site visits. Researcher Harry Collins asserts that personal interactions are important for transferring tacit knowledge; for example, direct interactions allow the more experienced scientists to physically demonstrate how they perform tasks, while less experienced scientists can use the opportunity to perform tasks under guidance and ask questions as needed. According to Collins, five types of tacit knowledge are transferred during a personal visit: concealed knowledge, mismatched salience, ostensive knowledge, unrecognized knowledge, and uncognized/uncognizable knowledge. Because knowledge recipients can ask knowledge sources questions during site visits, knowledge sources can better identify and articulate their concealed knowledge. Mismatched salience occurs when the knowledge source and recipient each focus on different important variables associated with difficult experiments, such that the knowledge source

---

81 Collins, “Tacit Knowledge, Trust, and the Q of Sapphire.”
82 Ibid.
83 Ibid.
fails to realize that the recipient needs to be told to do things a certain way and the recipient does not know the right questions to ask. Site visits resolve this problem because the knowledge source and recipient can watch each other work. The opportunity to watch a more experienced person work is also important for transferring ostensive knowledge—information that can be understood through direct pointing, demonstrating, and feeling—and helps knowledge recipients acquire muscle memory and interpret vague, sensory-related descriptions in protocols. Additionally, when the knowledge recipient emulates the knowledge source’s habits, unrecognized knowledge is transferred, though neither party recognizes that important knowledge has been transferred. In this way, knowledge is transferred through apprenticeship and emulation. Malefactors who wish to use synthetic genomics or synthetic biology to create harmful pathogens would have a greater chance of success if they had direct, physical access to individuals who have firsthand knowledge of the technologies.

Economic studies of the biotechnology industry affirm Collins’ and Vogel’s observations that practicing the routines of more experienced scientists is essential to transferring tacit knowledge. According to economists Lynne Zucker, Michael Darby, and Marilynn Brewer, it is important for small and new biotechnology firms to form alliances with academic scientists who have invented a new technique or technology because the inventing scientist can more easily transfer his or her tacit knowledge to recipients in the new company, a process that is important for translating research into
products. Zucker and colleagues demonstrate that research collaborations between scientists at biotechnology firms and university inventors have a significant positive effect on firm performance because the firms produced more patents. These findings further demonstrate the role that more experienced researchers play in overcoming contingencies in the laboratory. Malefactors who lack access to such researchers would likely have a more difficult time overcoming difficulties in using synthetic genomics or synthetic biology.

Several factors might influence the success of these direct interactions among scientists. One factor that facilitates the transfer of tacit knowledge involves the nature of the recipient’s existing knowledge base. If recipients can relate the new knowledge with their existing knowledge, they can absorb the new knowledge more readily. In addition, knowledge transfers are easier if recipients and donors share common knowledge, such as educational backgrounds or work experience. This suggests that malefactors who want to use synthetic genomics and synthetic biology to create a harmful pathogen will be more successful in acquiring knowledge if they share similar education backgrounds with any expert who may provide them with knowledge.

Another factor that facilitates the transfer of tacit knowledge is the knowledge recipients’ ability to trust that they will achieve positive results, even if they experience

86 Ben Ouagham-Gormley, Barriers to Bioweapons.
failures at first. For example, Collins describes a collaboration between British scientists and former Soviet scientists to measure the quality (Q) of sapphire, which was part of the laser development process. Soviet scientists had been able to obtain high Q measurements, which Western scientists had tried to replicate during the Cold War using the same methods to no avail. The method the Soviets used involved hanging the sapphire at the end of a thread that had animal grease on it. It was not until the Soviet Union dissolved that Western scientists learned what the Soviets had done: during a scientific exchange, a Russian scientist who had measured the Q of sapphire told a British scientist that instead of using animal fat to grease the thread on which the sapphire was suspended, he used oil that was behind his ears or under his nose. The British researchers adopted this technique but did not obtain high Q measurements until a year later. According to Collins, the British scientists struggled to achieve high Q measurements after implementing the Russian method of using human skin oil, and they eventually began to doubt that the Russians had ever actually achieved the high measurements; this distrust contributed to their yearlong delay in achieving high measurements. It was not until the British scientists witnessed the Russian scientists measuring Q that they understood that the high measurement could be achieved; in addition, they also saw that the Russians sometimes failed to achieve these measurements for no particular reason. This demonstrates the importance of trusting that an experiment will eventually succeed, as well as the crucial role that personal interaction plays in transferring knowledge—the

88 Collins, “Tacit Knowledge, Trust, and the Q of Sapphire.”
89 Ibid.
British scientists would never have known to use human skin oil, nor would they have seen firsthand that the method could succeed, if Russian scientists had not been physically present to share that knowledge and demonstrate the protocol. Malefactors who are gaining tacit knowledge from experienced scientists would need to learn that they will likely experience a number of failures before they succeed. Without this information, terrorists might abandon their work to search for solutions.

One overarching factor that facilitates knowledge transfer between individuals is their organizational environment. Research shows that organizations that stimulate employees’ free-thinking and creativity may be particularly valuable in dynamic environments like the biotechnology industry, where technological change is rapid and necessary. For example, research shows that new knowledge flows out of universities more readily than it does from commercial organizations, which is thought to stem from the creative intellectual environments fostered in universities compared to commercial firms. These intellectual environments are due in part to the social networks created by academic scientists, who exchange information freely without fearing the violation of their property rights, since these rights are protected through presentations and authorship of published research.

Another knowledge facilitator that relates to the organization in which individuals work is the organization’s so-called “transactional knowledge.” Transactional knowledge

---

90 Ibid.
92 Jaffè, Tratjenberg, and Henderson, “Geographic Localization of Knowledge.”
describes the organizational and management skills necessary to coordinate practices across multiple technical communities so that the knowledge of each community can contribute to the accomplishment of a larger technological goal.\textsuperscript{93} Research by Benjamin Sims on the interplay of tacit knowledge between nuclear designers and engineers shows that in order to reach compromises on design, members of both the design and engineering groups needed to speak and understand the other group’s language and judge whether the other group’s position was based on convenience or on actual technical necessity. In other words, the groups needed to be distinct but also needed to create a shared space so their respective knowledge could influence one another. These types of interactions are common in the biological sciences, where individuals with distinct but complementary skills work together towards a common goal.

While transferring knowledge between individuals can be effective in specific contexts, these interactions will be only partially effective at transferring knowledge for complex, interdisciplinary projects like synthetic biology research, in which different individuals or departments carry out different aspects of the project. Ben Ouagrham-Gormley identifies four reservoirs of tacit knowledge in scientific organizations such as those that conduct synthetic genomics and synthetic biology research: individuals, corporate culture, communities of practice, and organizational structure.\textsuperscript{94} These


\textsuperscript{94} Ben Ouagrham-Gormley, Barriers to Bioweapons, 36. According to Ben Ouagrham-Gormley, corporate culture is the product of an organization’s past experiences that is transferred to new employees via training, observation, or codes of conduct to convey the organization’s core values and practices. Since it relies on the organization’s past experience, it is a reservoir of tacit knowledge. Communities of practice
reservoirs are interdependent; having access to one is not sufficient to replicating past work. This is because, in complex projects, each reservoir carries out a particular aspect of the project and typically does not possess the same tacit knowledge that the other members involved in the project possess.\textsuperscript{95} Thus, no one reservoir contains all the tacit knowledge required to replicate the study. It follows, then, that malefactors would have a greater chance of acquiring knowledge if they seek information and tacit knowledge from multiple reservoirs.

Thus, the literature on tacit knowledge acquisition and transfer strongly suggests that although there are a number of factors that influence the success of knowledge transfer, malefactors would likely experience difficulties acquiring the knowledge necessary to use many biotechnologies if they did not have physical access to individuals with firsthand knowledge about the technologies they wish to exploit. Moreover, because of tacit knowledge’s local nature, malefactors who have learned how to create harmful pathogens using synthetic genomics and synthetic biology would need to adapt their tacit knowledge to the specific laboratory where they intend to create the pathogen. Lastly, because the majority of synthetic genomics and synthetic biology research is interdisciplinary in nature, involving many reservoirs of knowledge, access to only one of

\textsuperscript{95} Ibid., 28.
those reservoirs would limit the ability of a terrorist or a terrorist organization to replicate previous work.

In summary, STS literature demonstrates that because synthetic biology experiments are still novel, they are difficult to perform and their success hinges on trial and error and researchers’ tacit knowledge. However, no study has identified the type of knowledge involved in synthetic genomics, the contingencies associated with that technology, or the primary ways in which those difficulties are overcome. To determine how easy it would be for terrorists or other malefactors to synthesize genes using published protocols and commercially available technologies, I interviewed representatives of three U.S. gene synthesis firms to determine the extent to which they encounter technical difficulties. These firms are experts at producing gene-length DNA and have devoted considerable resources toward acquiring knowledge and optimizing the gene synthesis process so that it is as automated as possible. The following chapter provides the results of these interviews, including a discussion of how synthesis firms overcome contingencies. To determine how easy it might be for malefactors to assemble a genome from genes they ordered from gene synthesis firms, I also conducted a case study of the JCVI’s synthesis of the large, 1.08 Mb *M. mycoides* genome, which was assembled in a step-by-step process by combining small DNA fragments provided by gene synthesis firms into ever-larger fragments until the entire genome was assembled. These results are provided in chapter 5.
CHAPTER 4
DIFFICULTIES SYNTHESIZING GENES: INTERVIEWS WITH GENE SYNTHESIS FIRMS

Since the publication of research describing the de novo synthesis of pathogens like the poliovirus and the mousepox virus, gene synthesis has garnered much attention from the biodefense community, due to concerns that it might be diverted for nefarious purposes. There is, however, no empirical evidence to suggest that a malefactor could use gene synthesis to generate a pathogen. While previous research indicates there is a great deal of technical difficulty associated with synthetic biology, limited work has been published on the technical difficulties associated with gene synthesis. Yet, an understanding of the difficulties would provide valuable insight about how easily malefactors could synthesize pathogen genes. Current methods for gene synthesis can be difficult even for laboratory researchers to perform and reproduce, so they outsource synthesis to commercial suppliers who are experts in the process of gene synthesis.¹ In this chapter, I describe the difficulties that these expert commercial suppliers encounter when synthesizing genes, based on interviews I conducted with scientist-managers at three U.S. synthesis firms. My interviews indicate that there are, in fact, many contingencies associated with gene synthesis. Specifically, synthesis companies encounter difficulties four key phases of the process: during gene design at the beginning

¹ Marsic et al., “PCR-Based Gene Synthesis” (see chap.1, n. 14).
of the process, and with gene assembly, cloning, and verification at the end of the process. Difficulties are primarily overcome by consulting standard operating procedures (SOPs) and transferring problems to more experienced scientists within the companies who possess specialized skills not found elsewhere. These findings show that untrained terrorists or other malefactors would have a difficult time synthesizing genes themselves, indicating that they would also have a difficult time synthesizing pathogen genomes.

**Methodology**

In order to identify the technical difficulties associated with gene synthesis and the means by which these difficulties are overcome, I conducted interviews with four representatives from three gene synthesis companies from May to July of 2011, plus a follow-up interview with one of the companies in August 2013 to ensure the original interview information was still valid. All companies are headquartered in the U.S., and one company has facilities around the world. All interviewees were managers at their respective companies but had a scientific background and were therefore familiar with the process of gene synthesis. Due to the competitive environment of the gene synthesis industry, only three companies were willing to participate in my research. All companies are of similar size. One of the companies typically synthesizes short fragments of up to 50 kb (as of 2011), while the other two specialize in synthesizing larger sequences of up to 250 kb (as of 2015). Due to proprietary concerns, all interviewees asked to remain anonymous. Appendix E provides my initial set of questions for these interviews.
Overview of the Gene Synthesis Process

Gene-length DNA can be synthesized using a variety of methods. Synthesis firms, however, use proprietary methods to synthesize genes and, due to the highly competitive nature of the industry, do not disclose their methods. However, synthesis firms’ websites and scientific literature provide a broad overview of the process. In general, the process involves five steps: gene design and optimization, oligonucleotide synthesis, gene assembly, cloning, and verification. These steps are illustrated in figure 5 and described in the following sections.

Figure 5. Steps in the gene assembly process.

The success of the process hinges on customers ordering the correct genetic sequences. Many published genome sequences contain errors, which might impact how the synthesized gene functions in a customer’s experiments. Like many technologies, gene synthesis is an interdisciplinary process, requiring scientists with backgrounds in molecular biology, biochemistry, and genomics as well as individuals with expertise in computer programming.

---

2 Czar et al., “Gene Synthesis Demystified” (see chap. 2, n. 4); and Kosuri and Church, “Large-Scale de Novo DNA Synthesis” (see chap. 3, n. 8).
3 Epstein, “Challenges of Developing Synthetic Pathogens.”
Gene Design and Optimization

After a customer submits an order to have a specific gene synthesized, gene design is typically the first thing that occurs. During the design stage, sequences are analyzed by software to identify where problems might arise during synthesis. The software monitors a sequence’s ribonucleic acid (RNA) secondary structures, adenine/thymine (AT)- and GC-rich stretches, complex hairpins,\(^4\) sequence repeats, sequences that are known to be toxic to cloning vectors based on published research and proprietary solutions to previous challenges, and other factors that might complicate synthesis or impact the level of gene expression.\(^5\) While software does reduce inefficiencies during the synthesis stage, it does not guarantee that each step of the synthesis process will proceed flawlessly. These sequencers generate a chromatogram showing the results of sequencing, as well as a computer program’s interpretation of the sequencing data.\(^6\) However, the computer program’s interpretation of the data is often incorrect, requiring scientists to manually check the data against the chromatogram.\(^7\) As such, the genes that are submitted for synthesis may have incorrect sequences, and design software may not identify these errors before synthesis.

Software is also used to optimize gene sequences so that they best fulfill the customer’s goals. To do this, the software monitors factors that affect a gene’s function, such as its codon usage, generates optimized variants for the sequence, and selects the

---

\(^4\) Hairpins are structures that occur when two regions on the same DNA strand pair to form a double helix that ends in an unpaired loop. These structures can be unstable and affect synthesis.
\(^5\) Czar et al., “Gene Synthesis Demystified.”
\(^7\) Ibid.
variant that best suits customer requirements. Thus, design and optimization tools are meant to increase the efficiency of the synthesis process and improve the quality of the gene after synthesis.

Oligonucleotide Synthesis

After a gene is designed and optimized, its nucleotides are assembled into several oligonucleotides. This typically occurs in instruments called DNA synthesizers. There are several methods for synthesizing oligonucleotides, but the most common is solid-phase synthesis, a multistep process in which the chain of nucleotides that will eventually become the oligonucleotide is built on a solid bead, one nucleotide at a time, with washing steps in between. The process is almost entirely automated and typically synthetizes between 96 and 384 oligonucleotides simultaneously. In addition, commercial synthesis companies typically use a highly optimized process to build their oligonucleotides, which allows them to achieve high throughput. However, the speed with which these companies can synthesize genes is not indicative of an easily repeatable process in a nonautomated setting. In fact, research by Sanghvi indicates that

---

8 A codon is a three-base segment of DNA (a three-nucleotide segment); the bases and their order determine what protein will ultimately be translated. Since there are four nucleotides (adenine, thymine, guanine, and cytosine, or ATGC), there are 64 possible codons ($4^3$). There are only 20 genetically encoded amino acids, so most amino acids are coded for by multiple codons. Bacteria and mammalian cells use different codons, and mammalian genes often express poorly in bacteria and vice versa. This can pose a significant problem in subsequent steps of gene synthesis, since many cloning vectors are bacterial cells. Codon optimization software, which is offered as a free service by many synthesis companies for their customers to use, analyzes the sequence order and the customer’s chosen vector and determines which codons are optimal for that gene’s expression, then suggests the appropriate DNA sequence to meet those ends.

9 Sanghvi, “Assembly of Synthetic DNA,” 25–27 (see chap. 2, n. 4). According to Sanghvi the bead is a small, polymeric porous structure that is chemically inert during DNA synthesis and serves as solid support for oligonucleotide synthesis. In 2005, the most popular bead for small-scale synthesis was made from glass or silica.

10 Sanghvi, “Assembly of Synthetic DNA.”

11 Kosuri and Church, “Large-Scale de Novo DNA Synthesis.”
oligonucleotide synthesis would be highly inconvenient without the use of automated synthesizers because the synthesizers can produce thousands of sequences in parallel, whereas nonautomated techniques would produce fewer sequences.\textsuperscript{12} Further, Kosuri and Church contend that there are limits on the length of oligonucleotides that automated synthesizers can generate and error rates are high.\textsuperscript{13} In addition, the yield for each step of solid-phase synthesis must be high, especially for producing long oligonucleotides. However, the high error rates associated with longer oligonucleotides can reduce yield. For example, the yield for a long, 200-nucleotide (nt) fragment might be 99 percent but, due to the high number of errors in longer nucleotide fragments, the final yield might actually be only 13 percent. Additionally, problems in adding new nucleotides to the oligonucleotide chain can result in errors like single-base deletions.\textsuperscript{14}

**Gene Assembly**

In this phase, double-stranded DNA fragments of at least several hundred bases in length are assembled from oligonucleotides. Specific industrial assembly processes are proprietary and cannot be discerned, but the literature describes several protocols for assembly, including in vitro assembly methods like PCA, thermodynamically balanced inside out (TBIO) synthesis, and ligase chain reaction (LCR).\textsuperscript{15} PCA and TBIO use the enzyme DNA polymerase to assemble oligonucleotides, while LCR uses the enzyme

\textsuperscript{12} Sanghvi, “Assembly of Synthetic DNA.”
\textsuperscript{13} Kosuri and Church, “Large-Scale de Novo DNA Synthesis.”
\textsuperscript{14} Ibid. A single-base deletion is a type of mutation in which a single nucleotide base is deleted from DNA or RNA. These deletions alter the sequence of the oligonucleotide that is being synthesized.
\textsuperscript{15} Czar et al., “Gene Synthesis Demystified.”
ligase. These in vitro methods rely on the melting and annealing of oligonucleotides and are typically used for relatively short sequences of up to 1 kb.\textsuperscript{16} Some experts believe that gene synthesis companies likely employ modified, scaled-up variations of any of these methods, though it was reported in 2006 that most commercial companies use PCA.\textsuperscript{17} According to a handbook published in 2014 by the synthesis company GenScript, protocols exist to produce sequences up to approximately 750 kb using PCA, whereas LCR is limited to producing genes shorter than 2 kb.\textsuperscript{18}

Assembling these smaller genes into longer molecules currently involves cloning strategies. The genes are first cloned into a plasmid. After their sequences are confirmed to be correct, the genes are combined into larger sequences—a process that can require several steps, depending on how long the final gene will be. According to Czar and colleagues, the success of these cloning-based assembly methods decreases as the sequence length increases. In addition, these in vitro strategies eventually lose their viability, and it becomes necessary instead to use in vivo recombination methods in organisms such as \textit{Bacillus subtilis}, \textit{E. coli}, or yeast. This typically occurs when researchers must make more complex, megabase (Mb)-length DNA of more than 20 kb. Constructing these large pieces of DNA in vivo is typically a time-consuming and costly process that involves cloning small pieces of DNA and then sequencing them to weed out errors. The resulting error-free DNA segments are then assembled into the final, large

\textsuperscript{16} Ibid.
\textsuperscript{18} “Gene Synthesis Service” (see chap. 2, n. 5).
construct and the large construct is sequenced a final time to ensure its accuracy.\(^{19}\) The nonassembly aspects of cloning are described below.

**Cloning**

At this stage, the synthesized DNA fragment is cloned into a vector,\(^{20}\) which can be introduced into the intermediate host cells by transfection or transformation, so that the synthetic sequence can be confirmed as correct during the verification process (which is described more in the following section). According to the GenScript handbook, synthetic genes are often designed to include sites that facilitate cloning, such as restriction enzyme sites.\(^{21}\) Much of the success in the cloning stage is contingent upon the nature of the gene, as well as scientists’ understanding of the cloning vector and their level of experience with molecular biology techniques. Some companies clone DNA inserts into vectors using restriction enzyme cloning, in which enzymes are used to cut the DNA at specific sites, while other companies have developed and used proprietary cloning techniques to enhance production.

While cloning is a standard practice in molecular biology, it is not without difficulty. Some sequences can be toxic to host cells, and some repetitive sequences and/or domains that are AT- or GC-rich might be difficult to clone. There are also limitations in the capacities of certain vectors. According to GenScript, long genes are

---


\(^{20}\) Cloning vectors are DNA molecules that carry foreign DNA into a host cell. These include plasmids, phages, and bacterial artificial chromosomes.

often difficult for host cells to maintain and propagate because, being large, they consume too much of the cells’ energy.\textsuperscript{22} Additionally, \textit{E. coli}, a common bacterial host cell, will sometimes combine or mutate the synthetic gene.\textsuperscript{23} According to Integrated DNA Technologies, restriction enzyme cloning can unintentionally affect gene function because adding nonendogenous restriction sites to a gene construct can create a synthetic sequence that differs from the natural sequence.\textsuperscript{24} GenScript’s handbook indicates that some genes alter the physiology of their hosts, creating atypical conditions that may be difficult for researchers to accommodate or even identify.\textsuperscript{25}

In addition to verifying gene sequences to ensure that they do not contain errors, cloning can also be used to create very large DNA fragments in a process called in vitro ligation. During in vitro ligation, DNA strands are joined together (ligated) in a cell (in vitro). However, there are limits to the host cells in which DNA fragments are cloned. \textit{E. coli} can only accommodate and copy DNA that is approximately 300 kb in length, while yeast can accommodate and copy approximately 2 Mb of DNA.\textsuperscript{26} For example, a researcher who wishes to create a 400 kb DNA construct would need to clone the construct in yeast because \textit{E. coli} can only accommodate 300 kb.

\textbf{Verification}

Due to the inherent potential for errors in every stage of the gene synthesis process, all synthetic sequences must be verified before they are delivered to customers,
and sequences that contain mutations must be removed or corrected. Many of these mutations are the result of errors in oligonucleotide synthesis and the assembly process itself.

Verification typically involves excising genes from clones and sequencing them using Sanger sequencing—an expensive and time-consuming process that is difficult to automate. Since the process can be tedious and expensive, commercial firms use a variety of methods to reduce the number of clones required to find the perfect sequence. These include enzymatic techniques that filter out errors and next-generation sequencing techniques that screen sequences during the oligonucleotide synthesis or gene assembly stages to identify and use only those sequences that contain no errors. These methods have cut error rates and streamlined the verification process. Following verification, the gene is delivered to the customer, along with critical information about its sequence, such as a readout from the DNA sequencer that shows the gene’s sequence.

Problems with automatic sequencers can have a significant effect on research projects. One report noted that incorrect sequences are so commonplace that being able to detect and correct them is an essential skill in molecular biology. Indeed, the failure of a gene synthesis company to identify such errors would significantly impact its ability to verify that it had met its customer’s request.

---

27 “Gene Synthesis Service.”
28 Kosuri and Church, “Large-Scale de Novo DNA Synthesis.”
29 Ibid.
30 Sequencing data can be sent physically or electronically via secure logins.
Thus, though many national security analysts believe that the gene synthesis process is almost entirely automated—and could therefore be replicated by knowledgeable and well-equipped malefactors—the open literature suggests that there are limitations in the technologies used in the gene synthesis process. The literature also suggests that there are limitations in human knowledge about the behavior of genetic material and microorganisms. My research aims clarify the extent to which these limitations apply to gene synthesis firms, who devote significant resources to overcoming such limitations. The following sections provide the results of my interviews with synthesis firms. They describe the difficulties that synthesis firms encounter during gene synthesis, the means by which they overcome those difficulties, and the organizational factors that may have helped them resolve the difficulties.

**Contingencies in Gene Synthesis**

Interviews that I conducted between May and July of 2011 with representatives from three gene synthesis firms indicate that there are difficulties in four key phases of the process: during gene design at the beginning of the process, and with gene assembly, cloning, and verification at the end of the process. These difficulties are primarily due to the unpredictable nature of DNA. Although the process relies heavily on automation, machines and software cannot always respond appropriately to problems, requiring the intervention of an expert who possesses skills that are not easily found among employees throughout the company. While problems occurring upfront during gene design can have a ripple effect on the rest of the process, interviewees indicate that the most problematic
challenges occur at the end of the process during gene assembly and cloning. These
difficulties are discussed below.

**Gene Design and Optimization**

The design and optimization phases are key components of gene synthesis, but the
companies I interviewed have varied approaches to these phases. One company invests
heavily in research and development (R&D) for gene design, while the others offer a free
online optimization service for their customers to use. The companies that provide the
optimization service reported no difficulty or contingency in design or optimization
because the customers take on that duty when using the service. This does not mean that
the customers do not encounter problems in these phases; however, the companies did not
have any information about customers’ contingencies when I conducted my interviews in
2011.

The company that was investing in gene design R&D reported that there is some
difficulty in determining which sequence would yield the best results, primarily because
the behavior of all sequences is not known and, as such, the success of the synthesis
cannot be predicted:

**Scientist-manager at gene synthesis company A:** There’s still some
knowledge that goes into designing genes. . . . The guy who runs the gene
design program knows it the best, so everything that’s complicated ends
up on his table at some point. Everything that’s straightforward is handled
right at the front line by the salespeople. Then there’s the in-between stuff
that ends up being handled by one of the more experienced sales people or one of the underlings in the expression group.

**Shannon Fye-Marnien (SF):** So why is it that the guy who runs the gene design program gets the hard stuff?

**Scientist-manager:** He’s seen more stuff . . . and it’s science, so stuff that we’ve seen 100 times is easy—they have good predictions of how that stuff’s going to behave. But stuff they’ve seen 10 times, they have less accuracy in their predictions. . . . So you go from one extreme being a click of a button—straightforward stuff—to the other extreme, which is bioengineering on a very deep level.32

This exchange indicates that limitations in the understanding of biology affect the ability to predict how synthetic genes will behave during synthesis; more importantly, this exchange also indicates that when difficulties in gene design do occur, their resolution requires specialized knowledge and expertise that a few employees have. The inability to predict the behavior of genes during synthesis could produce genes that would be difficult to synthesize, making it more likely that the synthesis process would be problematic for terrorists and other untrained individuals.

**Oligonucleotide Synthesis**

All interviewees agreed that the process for chemically synthesizing oligonucleotides is automated and lacks technical difficulty, a finding that is consistent

---

with Sanghvi’s assertion that oligonucleotides are not difficult to synthesize.\textsuperscript{33} As mentioned previously, oligonucleotides are synthesized one nucleotide at a time until the entire oligonucleotide is built. This indicates that the technologies that once were manual and extremely tedious could become automated over time.

None of the interviewees reported contingencies related to the reagents, including those used to synthesize oligonucleotides. However, one interviewee noted that every time the company receives a new reagent lot from an existing manufacturer or changes manufacturers, the lot goes through quality control and is quantified and calibrated.\textsuperscript{34} This is a standard practice in laboratories to ensure consistency in results, indicating that differences in reagent lots can affect synthesis.\textsuperscript{35} This is consistent with Vogel’s interviews about the poliovirus synthesis, in which the participating scientists believed the inability to grow HeLa cells was due to differences in manufacturers’ fetal bovine serum.\textsuperscript{36}

**Gene Assembly**

Interviewees estimated that 90 percent of their companies’ gene synthesis orders are relatively easy to synthesize, while 10 percent of the orders are not easily synthesized and require human intervention. This intervention is typically needed in the gene assembly and cloning stages. According to my interviews, while many problems—such as the potential for sequences to be toxic to or unstable in *E. coli*, the propensity for

\begin{itemize}
  \item Sanghvi, “Assembly of Synthetic DNA.”
  \item Company A, interview, August 15, 2013.
  \item Vogel, *Phantom Menace*, 79 (see chap. 1, n. 11).
\end{itemize}
certain sequences to recombine, and the inability of some sequences to assemble by PCA
or other methods—can be predicted or prevented during the design stage of synthesis, the
unpredictable nature of DNA is such that design software is unable to identify all
potential problematic sequences. As a result, sequences that will cause problems later in
the gene synthesis process—during the assembly and cloning stages—can be produced.
Further, because scientists do not know the behavior of all DNA sequences, this
information cannot be incorporated into design software. Indeed, most interviewees
reported that the problems encountered during the assembly and cloning stages occur for
unknown reasons, indicating that current methods for anticipating problems that might
arise during synthesis are not always sufficient. Consequently, human expertise in
molecular biology and chemistry, among other disciplines, plays an integral role in
managing and overcoming these contingencies. Interviewees who faced such
contingencies indicated that they routinely transfer these problematic sequences to more
experienced scientists.37

Most interviewees agreed that contingencies are rare when synthesizing smaller
genes. However, all interviewees asserted that synthesizing larger genes is complex.
Although the companies have dedicated much R&D to improving their abilities to
synthesize larger genes, the process remains difficult. Much of this difficulty is due to the
nature of DNA, since larger stretches of it provide more opportunities for recombination
events, enzymatic reactions, and other errors. As one interviewee described it:

37 Company A, interview, May 20, 2011; and interview with employee 2 of gene synthesis company B,
June 15, 2011.
Scientist-manager at gene synthesis company A: The synthesis piece [process] is the same for small and large genes . . . but once you get to the point where you want to put the pieces together [assembly], it becomes tricky . . . Putting 2 kbs into four is straightforward; putting four into eight is straightforward. Then when you get to putting eights together to sixteens, it’s more tricky. Then when you get from sixteens to sixty-fours, it gets more complicated and requires handholding [careful attention] and expertise. Assembly is the labor-intensive part . . . because you could put things together, but at the end of the day you need to resequence the whole thing to verify that you got the right sequence—that you didn’t drop a single nucleotide somewhere along those hundred thousand base pairs.

SF: Why is it more difficult?

Scientist-manager: Because . . . the bigger the chunk of DNA is, the higher [the] likelihood that your primers won’t bind properly or will bind somewhere else; you’ll get sequences that are identical throughout so you may have recombination, and maybe you’ll get something that’s toxic to the carrier that’s on this piece of DNA. The only reason it’s more complicated is that there’s more and there’s more likelihood of stuff happening. And then there’s a combinatorial effect. So if you have 1 kb versus 1,000 kb, not only are there things that can interact with one gene, but there are thousands of things that can interact with a thousand
elements, so you get a combinatorial explosion, which needs to be managed.\(^{38}\)

Thus, DNA’s unpredictable nature can introduce many errors during the assembly process, and the risk of errors increases as the length of DNA increases. Additionally, sequence verification becomes more time-consuming as the length of DNA increases.\(^ {39}\)

Considering the challenges faced by experts in the field, malefactors who wish to use synthetic genes to produce a biological agent will likely encounter significant difficulties in the assembly stage, and later they will encounter some of these same difficulties when they try to verify cloned variants’ sequences.

The current demand for large-gene synthesis is low. One interviewee noted in 2013 that orders for sequences longer than 50 kb were rare, occurring approximately once a month.\(^ {40}\) All company representatives interviewed for this dissertation expected the demand to increase and, at the time of the interviews, were improving their protocols in preparation for a future potential increase. However, since the market for large-gene synthesis is currently still low, the companies are not building the technology in earnest.

Because synthesizing large genes is currently difficult for these companies, synthesizing large genomes is beyond their technical abilities. It is important to note, however, that three U.S. gene synthesis companies—Blue Heron, DNA2.0, and GeneArt—participated in the synthesis of the \textit{M. genitalium} genome, while only one of them—Blue Heron—participated in the synthesis of the \textit{M. mycoides} genome. \textit{M.}

\(^{38}\) Company A, interview, May 20, 2011.
\(^{39}\) Ibid.
\(^{40}\) Company A, interview, August 15, 2013.
*genitalium* and *M. mycoides* are bacteria with large genomes that were assembled by researchers at the JCVI.\(^{41}\) However, the synthesis firms only synthesized a library of small individual genes (approximately 1 kb in length), verified their sequences, and sent them to JCVI scientists for subsequent assembly. While synthesizing the genes lacked technical difficulty, these two projects signaled a major leap forward for the industrial process of gene synthesis because the companies performing the synthesis had to rapidly increase their throughput to accommodate the large amounts of DNA that the JCVI team ordered.\(^{42}\)

In addition to problems synthesizing large genes, my interviews reveal that synthesis firms also encounter problems working with automated tools. As noted previously, some companies use robotic tools throughout the synthesis process to increase throughput. While mainstream threat assessments claim that these robotic tools reduce the amount of human skill required to synthesize genes, my interviews reveal that, instead, they introduce a new set of problems, which require expertise that cannot be captured in a machine or software. These tools are controlled by software that tells the tool which tasks to perform and how to perform them. However, the software must first be programmed by an expert with knowledge of the tasks that are to be performed. Furthermore, every gene synthesis order is different and requires different tasks, all of which an expert must program into the software. Thus, while tasks may be automated by a machine or robot, they still depend heavily on expert human knowledge. As one

\(^{41}\) Gibson et al., “Complete Chemical Synthesis” (see chap. 3, n. 24); and Gibson et al., “Creation of a Bacterial Cell” (see chap. 2, n. 17).

\(^{42}\) Company B employee 2, interview.
interviewee described, “Everything has to be explicit because robots have no tacit knowledge. You have to tell them every single step . . . and since everything you’re making is different . . . every time the robot runs, it’s doing a different pattern of moves. And so you have a database that writes out that pattern every time based on the specific set of orders that are on that plate.”

Another interviewee characterized this human-derived knowledge as applying to not only robotics but to every part of gene synthesis: “Everything is human-based. Knowledge is human-based. Once you sort out the knowledge, then you can define this and standardize it and make it into an operating procedure. And that operating procedure might be a machine, software, and enzyme, or whatever. But someone needs to figure out what technology is needed, what it contains, and then do the experiment to either validate it or toss it away.” This shows that human knowledge is required for more than just programming software; it is also needed to discern which technologies best lend themselves to automation and to test how those technologies should be used. Thus, the very tools that make gene synthesis easier or automated still depend on human knowledge to achieve success, and a user without in-depth knowledge of gene synthesis will have great difficulty using such tools.

In addition to requiring human knowledge to function, robotic tools are limited in their ability to replicate human tasks, and scientists and programmers must modify protocols that are generally suitable for humans so that they can be both understandable

---

43 Company B employee 2, interview.
to and feasible for robotic technologies. This process of translating protocols so that they can be understood by robotic technologies also depends on human knowledge: an individual with knowledge about programming who can write the robot’s script, as well as modify it when protocols change. Programming knowledge is also needed to work around the limitations of the automated tools. As one interviewee explained, “There’s no fully automated gene synthesis, so . . . the protocols change all the time and people are continuously working around the database to get things done and shipped on time. . . . So . . . you’d say, “Well, we’re going to pretend that we’re doing this so it’ll flow through the database, but actually we’re doing that.”

Because of the tools’ limitations, scientists and programmers must sometimes work around the limitations by essentially tricking the tool into thinking it is dictating one task when, in reality, researchers are instructing it to carry out a different protocol. Thus, this suggests that, while databases and automated instrumentation may increase the overall efficiency of gene synthesis, they can also hinder the workflow when it is difficult to program human laboratory tasks or modify existing programs.

My interviews illustrate the vital role that the expertise of specific employees still plays in the synthesis process, despite the level of automation that has been achieved in the field. This knowledge is not widespread and is often tacit. As such, inexperienced individuals, such as terrorists or other malefactors, would have a difficult time programming and using automated systems; they would also have problems resolving difficult sequences in the laboratory.

---

45 Company B employee 2, interview.
Cloning

Interviews indicate that there is also some contingency in the cloning process. According to my interviews, much of the success in the cloning stage is contingent upon the nature of the gene, on scientists’ understanding of the cloning vector, and on scientists’ level of experience with molecular biology techniques. One interviewee characterized the process like this:

If the cloning sites are not chosen well by the researcher, the restriction enzymes alone can be complex. . . . When we go into a customer’s vector, unless we’re using a customer’s vector on a very regular basis, . . . we are using standard restriction enzyme cloning. The success of that depends on the stability of *E. coli*, as well as the design on the choice of the restriction enzymes. If you choose restriction enzymes together, cloning can be very difficult. If you do single-enzyme cloning, that can be very difficult. All those standard molecular biology hurdles [are involved].46

Another interviewee stated that in approximately 90 percent of the cases, it is easy to clone genes into vectors but sometimes, “for whatever reason,” the process is complicated.47 This suggests that in spite of all the advances in molecular biology and automation, some inherent challenges remain. This finding is consistent with Jordan and Lynch’s observations of PCR, a supposedly automated technology that still presents difficulties for researchers.48

---

46 Interview with employee 1 at gene synthesis company B, May 27, 2011.
47 Company A, interview, April 5, 2011.
48 Jordan and Lynch, “The Dissemination, Standardization, and Routinization” (see chap. 1, n. 11).
My interviews suggest that researchers—in addition to possessing familiarity with vectors and experience in molecular biology techniques—must also be patient, flexible, and capable of problem solving throughout the gene synthesis process. So, not only will terrorists and other malefactors encounter difficulties in assembling genes, but they will also experience difficulties when performing standard molecular biological tasks related to synthesis, such as cloning. These difficulties would likely affect the eventual success of any attempt at synthesis of a dangerous pathogen.

**Verification**

As mentioned above, errors in DNA sequence information are commonplace in the biological sciences due to limitations in automated DNA sequencers. Predictably, gene synthesis firms also experience problems interpreting information derived from these sequencers. These problems affect the companies’ ability to verify that the genes they synthesize have the correct sequence and are ready to deliver to the client.

Automated DNA sequencers report results using (1) a chromatogram that uses four colors to represent each nucleotide and (2) a text file of sequence data that is generated by a computer program. In theory, the results of the chromatogram can be interpreted easily. Figure 6 depicts an ideal, easy-to-read chromatogram with even spaces in between peaks. These spaces help scientists know how many nucleotides are in the sequence and in what order they appear.
However, interviews indicate that there is some ambiguity in interpreting chromatograms, which can confound sequence verification. One interviewee explained:

A typical problem that we see often is a run of nucleotides [a single nucleotide that appears consecutively in a sequence]. So if you have fifty As in a row, it ends up being very problematic to sequence. . . . It’s really hard to get sequence verification for fifty As in a row. . . . [A] DNA sequencer produces a chromatogram with peaks for every nucleotide. That’s fine as long as you have a distribution, but the moment you start getting runs on any given nucleotide, the peaks tend to flow into each other so it’s hard to tell if you have 50 or 51 peaks because it ends up being one big hill of peaks together. So that’s something we often see. . . . We probably synthesized the thing, but we can’t for the life of us sequence verify it. So depending on the customer and the order, you can deal with that by sprinkling in a single T [adding a thymine or other

49 “Interpretation of Sequencing Chromatograms.”
nucleotide to the sequence] in there every tenth or twentieth nucleotide, but that depends on the customer and what they’re trying to achieve.\textsuperscript{50}

As the interviewee noted, chromatograms can become difficult to read if a single nucleotide is repeated consecutively. In these cases, the nucleotides appear as a single wide peak, which can make it difficult to determine exactly how many nucleotides are in that peak. Figure 7 shows an example of a run of five adenines (As) in a row, making it difficult to verify the number of As that are actually in the sequence.

\textsuperscript{50} Company A, interview, May 20, 2011.
According to the interviewee, one way that synthesis companies can reduce the chromatograms’ ambiguity is to add another nucleotide to the sequence, such as a thymine (T). This would place a different-colored peak in the middle of the As, making it easier to determine how many As are in the sequence. Like other knowledge required for gene synthesis, this problem-solving knowledge is derived from individuals with specific skill sets. Therefore, successful verification of some sequences is contingent upon companies’ ability to work around the limitations of laboratory instrumentation while also meeting customer design needs.

51 “Interpretation of Sequencing Chromatograms.”
Summary of Difficulties

Interviews indicate that contingencies are encountered in the design, assembly, cloning, and verification stages of gene synthesis. Figure 8 summarizes the contingencies that can be encountered in all five stages.

![Gene synthesis process diagram](image)

Figure 8. Contingencies in the gene synthesis process.

As shown in Figure 8, challenges that require the involvement of experts with hands-on experience occur at key steps of the process: at the beginning during gene design, and near the end with gene assembly and cloning. Even the difficulties that are caused by equipment (verification) require specific expertise to resolve the issues. In other words, despite advances made in automation, synthesis is still a manual, expertise-based process; it is not the straightforward and fully automated process that is often described in threat assessments. Indeed, these difficulties persist despite the resources available to gene synthesis companies, which include experts who possess specialized knowledge in how to synthesize genes. As such, it is unlikely that terrorists or other malefactors—including those with technical ability and infrastructure—would be able to synthesize genes as easily as many national security analyses indicate.
Resolved Contingencies

Interviews conducted for this dissertation revealed that synthesis companies resolve difficulties by (1) consulting SOPs, including SOPs for how to deal with problem sequences; (2) transferring problem sequences to more experienced scientists; and (3) transferring knowledge by consulting published protocols, observing more experienced scientists, conducting training, and networking with other scientists. These methods are described in the following sections, and they indicate that gene synthesis firms possess a significant amount of in-house knowledge that has been custom-built and is not publicly available.

Consulting SOPs

The most common method for conveying knowledge and resolving difficulties within gene synthesis companies is by consulting SOPs. All companies interviewed use SOPs for every phase of the synthesis process, and they have additional SOPs for complex sequences. The companies’ experienced scientists, particularly those in the R&D departments, are the primary sources of SOP information; academic literature is a less common source of information. The companies frequently improve the synthesis process by modifying their SOPs to reflect new discoveries that R&D scientists have made, as well as new knowledge that employees obtain about how to synthesize a difficult sequence. The majority of SOP modifications are minor, such as buffer changes, and therefore do not require that scientists be retrained to acquire the new information. To the greatest extent possible, companies use all of the new knowledge they acquire to

---

52 Scientist-manager at company A, e-mail message to author, August 22, 2015.
optimize the synthesis process. A scientist-manager at gene synthesis company A explained how the company incorporates new information:

When we have a given project and we run into a problem because it’s something that recombines or something, you try to solve that particular issue [in the laboratory]. Sometimes you can solve things . . . if you bang your head on something long enough, by chance you manage to get through it. But also we try to capture that information [in an SOP] to try to figure what exactly went wrong and how we can prevent that in the future and then figure out tools, techniques, machines, [and] software to make sure we avoid it in the future—either avoid it by making sure we avoid it on the design side by making sure we can synthesize the gene while avoiding that particular variable, or by figuring out a way to synthesize or sequence it without that problem.  

Another interviewee described the process in similar terms:

SF: So someone encounters a problem, figures out how to solve it, and incorporates it into the SOP?

Scientist-manager at gene synthesis company B: Often [information about problem sequences is] incorporated into the design software. So the way the process flows is that an order comes in, and then it’s analyzed through design software. And what that did was to populate the database with a set of fragments to be built and a recipe for assembling those

---

together. Plus, we had a set of protocols [SOPs] that go with each of those kinds of fragments and so whenever you run across any depth . . . that design software embodies a whole bunch of rules based on having run across problems. Say you run across a problem, you see it a couple times, and it becomes a design rule. Then instead of having to loop back when you see that, it’s recognized ahead of time. Or things that you can’t recognize ahead of time, you have some way of recognizing once they happen. When we see it, we know what to do.\textsuperscript{54}

This indicates that when companies figure out how to synthesize difficult sequences, they inculcate that information into all of their SOPs, especially gene design software, so that they can prevent or overcome those challenges in the future. This process of incorporating new knowledge into SOPs and creating design rules allows companies to optimize the synthesis process by identifying problem sequences they have already encountered and using the knowledge their employees gained when solving the problem the first time to prevent (via gene design) or solve (via gene design or revised SOPs) the problem in subsequent orders. However, while SOPs are the primary means by which companies resolve difficulties, they only account for sequences that are relatively easy to synthesize, difficult sequences that the company has synthesized before, or difficult sequences that design software can work around. These SOPs are inadequate when companies encounter problematic sequences they have not synthesized before.

\textsuperscript{54} Company B employee 2, interview.
**Transferring Problem Sequences to More Experienced Researchers**

As shown in the previous section, companies’ first course of action when encountering a difficult sequence is to consult SOPs. These SOPs are based on the knowledge of companies’ more experienced scientists, particularly those in their R&D departments, and are modified whenever the company acquires new knowledge that optimizes the synthesis process. However, due to the unpredictable nature of DNA, these SOPs are often insufficient for helping companies synthesize difficult sequences. When SOPs are insufficient, companies rely on the expertise of more experienced researchers to resolve difficulties.

As described in previous sections, one company transferred complicated sequence orders to the individual who runs the gene design program because that individual had “seen more stuff.” Additionally, companies transfer difficult sequences to their R&D departments if their SOPs fail to work.\(^{55}\) One company reported that individuals in its production group—the group responsible for synthesis—generally have 10 or 15 years of laboratory benchtop experience, whereas scientists in its R&D group typically have doctoral degrees, 15 to 20 years of laboratory benchtop experience, and specialize in working on complex scientific problems. These qualifications indicate that individuals in the R&D group are more experienced and, ostensibly, more qualified to resolve contingencies.\(^{56}\) According to this interviewee, there are differing levels of expertise even within the production group:

---

\(^{55}\) Company A, interview, May 20, 2011.

\(^{56}\) Company B employee 2, interview.
There’s people with different levels of experience in the [company]. . . . There are a few experts within the production group who are better at figuring out which of the standard things [methods for resolving difficulties] to apply to a problem, and then once it reaches a certain point [of difficulty], it gets tossed over the fence to the R&D guys as a source of opportunity to them. . . . So you have a mix of people [in the production and R&D groups] who have more or less experience.57

Much of the expertise that R&D groups and other experienced scientists in synthesis companies have is the result of years of trial and error. Interviews indicate that trial and error is a somewhat unavoidable process in gene synthesis, and that the information that results from trial and error is incorporated into SOPs to improve the synthesis process. However, despite the fact that trial and error appears to play some role in resolving difficulties with complex sequences, or in finding new methods to optimize the synthesis process, two interviewees were not comfortable with saying their companies used trial and error. One interviewee was uncomfortable with the term “trial and error” itself because it seemed to imply that scientists were engaged in shoddy or random scientific work:

I wouldn’t call it trial and error; I would call it systematic problem solving. Trial and error, to me, is just sort of a [way of saying] “randomly pick anything and do it.” But systematic problem solving is a very defined process. You have a problem, you sort which are the variables, you test all

57 Company B employee 2, interview.
the variables in a systematic fashion, . . . and then you figure out why that variable is a problem and then you fix that.\textsuperscript{58}

While uncomfortable with the term “trial and error,” this interviewee regarded the process of systematic problem solving as an essential component of performing proper scientific work. However, the interviewee noted that systematic problem solving is rarely done because 90 percent of orders are easily synthesized and do not require problem solving, so companies are reluctant to spend time solving issues with difficult genes.\textsuperscript{59}

Another interviewee was uncomfortable with the concept of trial and error from a business standpoint:

\textbf{SF:} How do you handle the unexpected technical problems that come up?

\textbf{Scientist-manager at gene synthesis company B:} There’s a series of standard procedures. . . . Typically at every step along the way, there are two or three different ways to accomplish that step, and the method is chosen based on economics and some analysis of the sequence. So the things that are hard and expensive to do work on almost everything, and the things that are easy to do work on most things, so you want to optimally parse moving parts through the process so that only things that need to go through the expensive methods do. And then when you get certain classes of failure there’s another set of steps you take. So there’s standard procedures [SOPs] for dealing with the exceptions [difficult sequences].

\textsuperscript{58} Company A, interview, April 5, 2011.
\textsuperscript{59} Ibid.
So, for instance, if the gene codes a protein that’s toxic in *E. coli*, then you put it into a different vector, or you grow the cells differently, or you do a different strain. There’s a whole lot of ways to solve that. But you don’t ever want there to be a trial and error process involved; you want to have a set of procedures that lead to at some point saying [to the customer], “We probably can’t do this without spending a lot of money on it. Do you want this to be a molecular biology project, or should we get back to you?”another interviewee from that same company affirmed that the company did not engage in trial and error: “There’s no trial and error—we have protocols for those [nonautomated processes]. . . . It’s just there are some nonautomated processes. I wouldn’t call it trial and error because that assumes that you don’t know how something’s going to work. But 97 percent of the time it works, the protocol works; it’s just that when you run into a molecular biology issue, you have to make changes to your protocols.”

Interviews reveal that while synthesis companies rely on the knowledge of their more experienced scientists, such as those in R&D groups, they avoid using trial and error methods to solve difficulties because they are too expensive and time-consuming. In the few cases when trial and error is required, synthesis companies typically consult with their customers first to determine whether the customer is willing to pay the resultant costs. Still, due to the unpredictable nature of DNA, the efforts of companies’ more experienced scientists might still fail to result in a successful synthesis.

---

60 Company B employee 2, interview.
61 Company B employee 1, interview.
Transferring Knowledge

Synthesis companies may also attempt to resolve difficulties by transferring knowledge within the company, although this method is less commonly used. To transfer knowledge, companies will use academic literature or published protocols; conduct training, which involves junior scientists replicating the routines of more experienced scientists; and network with other scientists.

From literature and published protocols, companies can glean new knowledge about sequences and their behavior; techniques that improve the synthesis process; and, less commonly, how to resolve difficulties in synthesizing complex or large sequences. Still, most interviewees reported that the majority of changes their companies make to SOPs derive from internal sources—specifically, from R&D departments’ efforts to optimize the synthesis process. Interviews indicate that these more experienced scientists do consult literature when attempting to fill complex orders that contain requirements with which they are not familiar, but they must still adapt the knowledge gained from the literature to their specific needs. An employee of company A explained the role of literature in the gene design process:

So, say someone just wants a standard gene expressed in E. coli. Fine, boom, they [the company’s scientists] can do that; typically that’s easy. Then it could be an order from someone who already has something that expresses really well in E. coli and they want to express it even better, plus they have some different constraints, some measures. Then there’s a combination of trying to look at the literature—it’s like any other science—
what has been done on similar things in the past. And then you have to propose ten different experiments to sort out what variables can you adjust according to X, Y, and Z. How do you design that experiment, what variables do you want to test, how do you test them systematically? And you design the experiments, you build the ten genes, test them in the system, analyze the data, come to conclusion, and you predict based on these new data points what the next gene should be, etc.\textsuperscript{62}

This example shows that the literature and other published protocols inform the synthesis process, especially when companies are attempting new tasks. However, the process that the interviewee describes above suggests that gene synthesis scientists must determine how to use information found in the publications to test experimental variables. This indicates that while published literature can serve as a starting point for helping scientists resolve difficulties, it does not provide information about how to carry out the experiments. Instead, gene synthesis scientists must interpret the information in the literature and adapt it to their specific needs. After doing this, the scientists can then test the new information in the laboratory using trial and error. When the resulting information is instrumental in resolving a difficulty or optimizing a process, scientists incorporate it into their SOPs.

In addition to using SOPs and consulting academic publications, gene synthesis companies also conduct training to transfer knowledge from more experienced scientists to less experienced ones. All companies reported that their employees undergo training

\textsuperscript{62} Company A, interview, April 5, 2011.
that involves observing more experienced scientists and replicating routines. Typically, most companies cross-train their employees so they are capable of performing functions at numerous points in the synthesis process. One interviewee explained that new staff typically spend the first three months attending training sessions and observing more senior scientists in their work. Once these employees have demonstrated that they can reproduce results reliably, they are allowed to work more independently. An employee of company A explained, “They [new employees] are trained by someone . . . who has been trained before on that particular instrument or that particular process. So they’re trained on that system. And first they just stand by and watch someone else do it, and then they do it themselves while having someone watch over them, and then eventually they’ll run with it . . . follow the protocol.” One company estimated that it could take weeks to months for a new employee to learn the process well enough to work independently, while another estimated that it could take up to six months for untrained individuals to work completely independently.

Thus, following SOPs alone is not sufficient for new employees to learn how to synthesize genes. Rather, new employees need hands-on demonstrations and training with an experienced scientist to learn how to use SOPs. This suggests that, in addition to not being able to rely solely on technology or formal education, a terrorist or other malefactor would not be able to rely solely on published protocol to synthesize genes and would instead need training by an experienced scientist in the field or task of interest.

---

63 Interview with scientist-manager at gene synthesis company C, June 8, 2011.
64 Company A, interview, April 5, 2011.
65 Company B employee 1, interview; and Company C, interview.
Synthesis companies can also obtain information that is pertinent to resolving difficulties from their customers, especially during the design and cloning stages, although this is a very uncommon method for resolving difficulties. In some cases, customers have information about the nature of a sequence that might be pertinent to the synthesis process, such as information about a sequence’s toxicity in a particular vector. In other cases, synthesis companies confer with customers about ways to work around complexities. An employee of gene synthesis company B shared an example: “They certainly run into complexities on the synthesis side [of the laboratory], . . . then there are sequences that are not stable in *E. coli* to propagate, so there’s a level of complexity there where decisions have to be made, and they’re made with the customer, such as moving to a low-copy vector to try to keep it stable in *E. coli* and prevent recombination.” This example indicates that synthesis companies sometimes consult with customers when they encounter difficult sequences, allowing customers to suggest which course of action to pursue. In other words, in spite of their expertise, gene synthesis scientists may not possess the appropriate knowledge related to certain tasks or even certain vectors and must, therefore, call on the knowledge of other scientists—in this case, their clients—to select the best approach for a successful synthesis.

In addition to consulting with clients, synthesis firms also network with other scientists to acquire knowledge. One interviewee emphasized the role of scientific conferences in transferring knowledge, noting that the presentations and networking opportunities can provide valuable insight about new techniques that could help resolve

---

66 Company B employee 1, interview.
contingencies or optimize the synthesis process.\textsuperscript{67} At least one company placed a strong emphasis on networking as a means to acquire knowledge.\textsuperscript{68} Another company described all of its employees as “science geeks” who like to talk to other scientists.\textsuperscript{69} However, these networking relationships take place outside the gene synthesis community of practice because competition within the industry is fierce and employees must take care not to disclose proprietary information. Instead, synthesis firms maintain their connections to the gene synthesis community of practice through industry groups like the International Gene Synthesis Consortium (IGSC). Although IGSC members refrain from sharing technical information that would reveal proprietary information, they do discuss the direction of the synthesis industry in general and collaborate with each other on common issues, such as strategies for maintaining biosecurity.

The methods that gene synthesis companies use to resolve technical difficulties are summarized in table 4. These are consistent with many of the knowledge transfer methods described in chapter 3, including replicating routines (observing more experienced scientists, conducting training) and consulting explicit knowledge sources (literature and published protocols and SOPs).

\textsuperscript{67} Company C, interview.
\textsuperscript{68} Company B employee 1, interview.
\textsuperscript{69} Company A, interview, May 20, 2011.
Table 4. Gene synthesis company methods to resolve contingencies.

<table>
<thead>
<tr>
<th>Contingency resolution method</th>
<th>Company A</th>
<th>Company B</th>
<th>Company C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Following SOPs, including SOPs for problematic sequences</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Transferring problem sequences to more experienced scientists</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Transferring knowledge (consulting published protocols, observing more experienced scientists, conducting training, networking with other scientists)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

All three companies used the same methods to resolve contingencies. Most of the methods for resolving difficulties hinged on the in-house knowledge resident in the synthesis firms, including (1) information used to generate and modify SOPs, (2) information about how to program design rules into design software and tasks into automation software, (3) information about how to synthesize some difficult sequences, and (4) information required to train new employees. Companies’ more experienced scientists possess much of this knowledge, and, due to in part to competition in the industry, the knowledge is not publicly available.

**What the Interviews Tell Us**

Interviews indicate that much of the work in commercial gene synthesis is automated and based on routinized practices. The companies primarily rely on SOPs to resolve difficulties. While SOPs are an example of explicit knowledge, the SOPs in synthesis companies are derived from in-house knowledge that is not publicly available. Moreover, the SOPs contain an element of tacit knowledge, as evidenced by the fact that new employees are required to observe the routines of more experienced scientists in order to learn about gene synthesis—not simply follow SOPs. Other forms of explicit...
knowledge that the companies use include gene design software and automation software that controlled robotic technologies.

Companies use academic literature primarily to obtain knowledge about sequences and their behavior and find methods for synthesizing large sequences. They also use the literature to discover methods that might optimize the synthesis process. However, the companies must then test these methods using trial and error and adapt them to their specific needs. In other words, the firms must customize the protocols to meet the needs of their working environment, a finding that is consistent with previous STS literature.\(^7^0\)

Although much of the work in commercial synthesis is automated and based on routinized procedures, the field is not completely automated and some of the work—including gene design, synthesis of complex sequences, large-gene assembly, and cloning—still requires human judgment and human skill. This judgment is needed to carry out tasks that are not automated or are only semiautomated; to know how to update SOPs, which are widely used; to amend design and automation programs; and, less commonly, to resolve contingencies in synthesizing complex or large sequences.

Interviews show that much of the judgment is based on tacit knowledge that is held by individuals—for example, a staff member who oversees gene design or scientists in the R&D department. Part of this knowledge is then made explicit when it is codified into SOPs or design rules. It is important to underscore that almost every aspect of gene synthesis depends on human knowledge, some of which is tacit, including knowing

\(^{70}\) Ben Ouagrham-Gormley, *Barriers to Bioweapons*, 27.
which process do and do not lend themselves to automation. In gene synthesis
companies, the tacit knowledge that is made explicit in the form of SOPs and design rules
seems to be held locally due to proprietary concerns. That is, those outside of the
synthesis companies appear to lack access to this kind of knowledge or at least lack
access to the breadth of knowledge that synthesis companies possess as experts in gene
synthesis.

The companies’ emphasis on training also points to the role of tacit knowledge in
successfully conducting gene synthesis, since much of the training involves observing
more experienced scientists at work. Additionally, one interviewee revealed that tacit
knowledge in the form of “lab hands”—the ability to work with equipment in an intuitive
and comfortable manner—might have a role to play in gene synthesis: “In science, you
see a lot of people who are highly skilled but have hands that don’t work right. . . . You
can be as smart as you want, but if you can’t pipette properly [you won’t be
successful]. . . . It’s not a Ph.D. skill that you need [to be successful in the laboratory].

In addition, my interviews reveal that much of the knowledge that gene synthesis
firm employees possess is specialized. Although interviewees reported that employees
are cross-trained so that they are adept at numerous aspects of gene synthesis, my
interviews show that some employees are more adept at gene design than others, while
other employees are more adept at synthesizing difficult sequences.

---

Interviewees also reported some difficulties with instrumentation and tools, such as reading DNA sequencers. Resolving these difficulties also involved tacit knowledge, in that scientists’ experience helped them interpret chromatograms.

**Implications to National Security**

Although commercial gene synthesis is highly automated and routinized, some of the work is not automated and requires human judgment and skill, and expert commercial firms still encounter contingencies when trying to synthesize complex or large sequences. Malefactors who attempt to create a pathogen genome using synthetic genes as starting material would also encounter difficulties. Because malefactors would lack access to the skill sets and other resources that synthesis firms possess, such as support from more experienced scientists, protocols that leverage extensive in-house knowledge, and training in how to use protocols, they would be less likely to overcome technical difficulties during gene synthesis. This dearth of resources is exacerbated by the gene synthesis industry’s highly secretive nature, in which companies do not make their methods known to the public; thus, terrorists will not be able to exploit any of their vast resources.

Consulting the literature may help malefactors overcome difficulties, but they would first need the knowledge to recognize which information in the literature is useful for their goals; without this knowledge, the literature would not be very helpful. They would also need to engage in trial and error to adapt the information to their specific needs. Furthermore, the literature may lack information on the pathogen or type of
sequence the malefactor is synthesizing, or it may provide the wrong sequence information altogether.

Although synthesis companies occasionally use trial and error to resolve difficulties, they are reluctant to do so due to time and cost considerations. Instead, they specialize in synthesizing sequences that protocols and more experienced scientists can handle, including smaller genes and genomes. These sequences make up the bulk of the orders the companies process. Academic or specialized laboratories currently specialize in synthesizing more complex sequences, such as longer genes and large genomes. If a malefactor wishes to synthesize these larger sequences to create a pathogen, a more efficient option would be to order smaller DNA sequences from synthesis firms, then assemble those smaller sequences into an entire genome using protocols described in the literature, and transplant the genome into a cell. However tempting this route might be for malefactors, creating an entire genome using DNA sequences obtained from gene synthesis companies is not an easy task, as will be demonstrated in the next chapter, which analyzes JCVI’s assembly of the *M. mycoides* genome and its subsequent transplantation into a cell.
My research findings described in the previous chapter show that scientists experience a number of technical difficulties when synthesizing genes on an industrial level, despite using automated tools in the synthesis process. Specifically, my research illustrates that malefactors who are attempting to synthesize pathogen genes using benchtop DNA synthesizers would have trouble synthesizing problematic sequences; producing longer genes; cloning certain sequences; and verifying that the gene(s) they synthesized had the correct sequence. To determine whether terrorists might encounter similar difficulties if they opted to order genes from synthesis firms and assemble those genes into a pathogen genome, I interviewed scientists at the JCVI who helped assemble and transplant the large, 1.08 million base pair *Mycoplasma mycoides* genome. The *M. mycoides* genome was built using smaller, commercially provided genes as starting material; the experiment therefore serves as a good proxy for determining what sorts of difficulties a terrorist might encounter in assembling a genome using commercially produced genes. After JCVI researchers built the *M. mycoides* genome, they transplanted it into a recipient cell, and the recipient cell’s genome transformed into *M. mycoides*; this experiment proved the hypothesis that researchers are able to create synthetic genomes.
and transplant them into cells, a process that terrorists would need to use if they were to create biological agents synthetically.

My interviews with JCVI scientists show that there is great difficulty involved in creating organisms when using synthetic genes as a starting point. These difficulties include assembling large genes and purifying them from yeast, transplanting genomes into host cells, using PCR, and verifying that gene sequences are correct. The interviews also show that, similar to gene synthesis companies, the JCVI’s scientists were able to overcome difficulties by using the knowledge of more experienced scientists. However, while gene synthesis firms prefer to synthesize shorter and predictable sequences so they can avoid costly and time-consuming trial and error, the JCVI relied heavily on trial and error to assemble and transplant the *M. mycoides* genome as well as to overcome other difficulties commonly encountered in their laboratory. The JCVI scientists often used published protocols as a starting point to solve problems, but these protocols had limited utility due to the novelty of the *M. mycoides* experiment. As a result, the scientists were required to create new methods and protocols throughout the course of the project. Considering that the JCVI team included dozen of experts and operated with all the required financial and material resources, the difficulties they encountered testify to the fact that genome synthesis is hardly a cheap and straightforward process. It follows, then, that terrorists who wish to assemble a pathogen genome would face even greater challenges, especially if they do not possess the appropriate expertise and resources and if the genome is large and complex.
**Methodology**

I interviewed one JCVI researcher who participated in the *M. genitalium* project, the results of which were published in 2008, and two researchers who participated in the assembly of the *M. mycoides* genome, which was published in 2010. The former researcher asked to remain anonymous and will be described as “anonymous JCVI researcher” in the citations. The latter two researchers are Dr. John Glass and Dr. Ray Chuang. All interviewees were doctoral-level researchers with at least 15 years of laboratory experience. Glass is a principal investigator, while Chuang and the anonymous researcher were directly involved in the day-to-day laboratory work. Chuang and the anonymous researcher participated in assembling and purifying the *M. mycoides* genome, while Glass oversaw efforts to transplant it into a host cell. While I was unable to interview Dr. Dan Gibson, the JCVI researcher who was responsible for assembling the *M. mycoides* genome, I reviewed online lectures and documentaries featuring him to discover the difficulties that he encountered during the assembly process.

My questions for the interviewees focused on difficulties they encountered and methods they used to resolve the issues. I conducted two of the three JCVI interviews over the telephone and the third interview at the JCVI’s Rockville, Maryland, campus. I recorded all the interviews with the interviewees’ permission.

**Background on the *M. mycoides* Project**

In 2010, researchers at the JCVI made history by announcing that they had created the first self-replicating synthetic bacterial cell. This project was part of the JCVI’s goal to create a so-called minimal cell, one that contains only the machinery
necessary to sustain its life.¹ In a 1999 paper, JCVI researchers asserted that one of the key steps in identifying the minimal gene set necessary for a self-replicating minimal organism would be to create an artificial chromosome.² To achieve this, the JCVI began attempting to create a minimal cell from *Mycoplasma genitalium*, a bacterium with the smallest known genome that can be grown in pure culture.³ Although *M. genitalium*—at 583,000 bp—is currently the smallest known bacterial genome, it is a large genome in comparison to viral genomes and other genomes that have previously been synthesized, such as the 5,300 nt phiX bacteriophage.

To synthesize such a large genome, JCVI researchers had to develop novel tools and methods for building large DNA fragments.⁴ The experimental method involved asking three gene synthesis firms to produce 101 gene cassettes that were between 5,000 to 7,000 bp long⁵ and then assembling those cassettes into the entire *M. genitalium* genome in five stages.⁶ Although the researchers succeeded in assembling the *M. genitalium* genome, Vogel’s research indicates that it was a tedious process.⁷ This was the first bacterial genome that had ever been synthesized and, at the time, the largest chemically derived structure created in a laboratory. The project stalled because

---

³ Gibson et al., “Complete Chemical Synthesis” (see chap. 3, n. 24).
⁴ Gibson, e-mail message to author, September 29, 2015.
⁵ John Glass, interview by author, March 4, 2013. According to Glass, JCVI researchers synthesized the first few 5,000 bp gene cassettes themselves but realized they lacked the manpower to synthesize all 101, so they outsourced the synthesis to Blue Heron, DNA2.0, and GeneArt.
⁶ Gibson et al., “Complete Chemical Synthesis.”
⁷ Vogel, *Phantom Menace*, 99 (see chap. 1, n. 11).
researchers could not transplant the genome into recipient cells; still, the project was the first successful attempt to assemble a large synthetic genome. At the same time they were working on *M. genitalium*, JCVI researchers were working with the *M. mycoides* bacterium and discovered that they could transplant its 1.08 Mb genome into *M. capricolum*, a closely related species.\(^8\) Once the researchers discovered that the *M. mycoides* genome could be transplanted into a recipient cell, they decided to create a synthetic cell using *M. mycoides* instead of *M. genitalium*, so they changed their focus toward that goal.

The *M. mycoides* project was the result of nearly 15 years of effort that required the JCVI team not only to develop new tools and techniques to assemble large segments of DNA, but also to develop techniques for transplanting synthetic genomes from one species into another.\(^9\) The project cost approximately $40 million and involved dozens of scientists.\(^10\) Widely hailed as a revolutionary project in the life sciences, the experiment proved that new self-replicating cells could be made using genes that were designed on a computer and produced with chemicals, rather than extracted from an organism.

While previous studies have characterized the technical difficulties in creating synthetic cells with smaller genomes, no study has examined the difficulty in synthesizing large bacterial genomes. My interviews with JCVI researchers reveal that they encountered a number of difficulties during the *M. mycoides* project, primarily with

---


\(^9\) Gibson et al., “Creation of a Bacterial Cell” (see chap. 2, n. 17).

sequence verification, purifying large genes from yeast cells, transplanting the assembled *M. mycoides* genome into the recipient *M. capricolum* cell, and even using common laboratory techniques such as PCR. The following sections describe the *M. mycoides* project methodology and the results of my interviews with JCVI researchers.

**How *M. mycoides* Was Assembled and Transplanted**

The *M. mycoides* project involved two stages: genome assembly and genome transplantation. During the assembly stage, the 1.08 Mb *M. mycoides* genome was assembled in multiple steps, which are shown in figure 9. The assembly process began with JCVI researchers designing the genome so that it could include watermark sequences,\(^{11}\) antibiotic resistance genes, and other sequences that would help them differentiate the finished, synthetic genome from a natural one. A key factor in the success of the genome’s design—and, indeed, the entire project—was the accuracy of the sequences the researchers used.\(^{12}\)

Next, Blue Heron, a gene synthesis firm, synthesized 1,078 gene cassettes that were each 1,080 bp in length (represented by the orange arrows in figure 9). These 1,078 cassettes were then recombined in yeast to produce 10,080 bp assemblies (represented by the blue arrows in the center of figure 9).\(^{13}\) The cassettes were then transferred to *E. coli*, because *E. coli* is better than yeast at propagating large

---

\(^{11}\) Watermarks are DNA sequences that spell out words or phrases that the gene designer wishes to convey, using the “alphabet” of genes and proteins. The *M. mycoides* genome’s watermarks include the names of the 46 researchers who helped with the project, as well as three quotations. Watermarks are useful for helping prove that genomes are synthetic and not native, and they can be used to identify the laboratory from which the organism originated.

\(^{12}\) Gibson et al., “Creation of a Bacterial Cell.”

\(^{13}\) Two of these 1,078 cassettes were pieced together in vitro using enzymatic methods, while the rest of the 1,076 cassettes were made using in vivo homologous recombination in yeast.
amounts (for example, 1,078 cassettes) of DNA. JCVI researchers isolated plasmid DNA from the individual *E. coli* clones and screened it to identify cells that contained the 10,080 bp assembled fragment. Cells that contained the 10,080 bp assemblages were sequence-verified to ensure they contained the correct sequence. Those that did were used in the next stage of assembly.

Next, the JCVI researchers used the same methods to combine the 109 10,080 bp assemblages into 100,000 bp assemblies; there were eleven of these (represented by the green arrows in figure 9). However, these larger assemblies were unstable in *E. coli* because *E. coli* was incapable of propagating sequences larger than 10,000 bp and because AT-rich genomes, such the *M. mycoides* genome, tend to be toxic in *E. coli*. Therefore, the recombined DNA was extracted from yeast cells, not *E. coli* cells. This required that the JCVI team improve their ability to extract 100,000 bp pieces of DNA out of yeast.

After extracting the 100,000 bp assemblies from yeast, they put all eleven back into a yeast cell so that they could assemble into the entire, 1.08 million base pair genome (represented by the red outer circle in figure 9). The scientists then performed PCR of the eleven junctions (where the eleven 100,000 bp fragments joined) to make sure they lined up correctly, and ensured the eleven sequences were assembled in the correct order. After confirming that the genome’s sequence was correct, the researchers then moved on

---

14 Gibson, “Assembling and Engineering.”
15 A plasmid is a small, circular, double-stranded DNA molecule that is distinct from a cell’s chromosomal DNA.
16 Ibid.
17 Ibid.
to the transplantation stage, in which they transplanted the genome into a *M. capricolum* cell. After transplantation, they sequenced the *M. capricolum* cell that had received the synthetic *M. mycoides* genome to ensure that it had received it. This confirmed that the synthetic transplant’s genome was identical to the synthetic *M. mycoides* genome that JCVI researchers had designed, including watermarks and antibiotic resistance genes. Ultimately the researchers succeeded in transforming the *M. capricolum* cell into *M. mycoides*.

The experiment proved that researchers could use a sequence from a computer, make some design changes, assemble an organism’s genome in yeast, and transplant it into a recipient cell. The experiment also provided the JCVI with a functional assay for creating a minimal genome: researchers can identify and delete nonessential genes from the synthetic *M. mycoides* genome to reduce it to only those genes that are essential for the synthetic bacterium’s survival.
Contingencies in the *M. mycoides* Experiment

Although the experiment was successful, JCVI researchers encountered a number of unexpected difficulties. For example, they were given an incorrect sequence from a synthesis firm, which required extensive efforts to identify. They also faced steep challenges during key stages of the process: when purifying large genes from yeast cells, when transplanting the assembled *M. mycoides* genome into its recipient cell, and even

---

when using common laboratory tools and techniques. These contingencies are described in detail in the sections below.

**Assembling and Purifying Large Genes**

Interviews with JCVI researchers echo synthesis companies’ assertion that large pieces of DNA are generally difficult to assemble. A JCVI researcher who helped assemble the *M. genitalium* genome described an experience synthesizing a 75,000 bp sequence for another project. The researcher was using PCR to assemble the 75,000 bp fragment, but the piece would not come together. The researcher extended the PCR annealing step, which normally takes 30 seconds, to last several hours and changed the temperature of the PCR reaction in several steps. The fragment eventually assembled, but the process was very time-consuming.\(^{19}\) The researcher suggested that some particular areas of genes or genomes are simply difficult to assemble and that this difficulty depends on the sequence—a finding that is consistent with what gene synthesis companies reported in my interviews with them. While the researcher was not aware of a similar problem occurring during the *M. mycoides* assembly, the experience confirms that assembling large pieces of DNA is generally a difficult process.

Assembling the *M. mycoides* genome was the responsibility of Dr. Dan Gibson, a molecular biologist who had been working at JCVI since 2004 and had previous experience assembling the *M. genitalium* genome. While Gibson had help assembling the *M. genitalium* genome, he assembled the *M. mycoides* genome by himself.\(^{20}\) According to Gibson, the *M. mycoides* genome assembly “went pretty quickly” because he had already

\(^{19}\) Anonymous JCVI researcher, interview by author, September 26, 2013.

\(^{20}\) Ibid.
developed methodologies for building large DNA fragments during the *M. genitalium* assembly, which was more difficult because he and his colleagues had to develop those novel methodologies from scratch.\(^{21}\) Dr. John Glass, who led the project’s effort to transplant the synthetic *M. mycoides* into a recipient cell, confirmed that assembling the *M. mycoides* genome was relatively straightforward.\(^{22}\)

Despite the relative ease of the *M. mycoides* genome assembly, Gibson and colleagues reported that they encountered problems propagating sequences larger than 10,000 bp in *E. coli*, indicating that the host cells they used during the project were unpredictable.\(^{23}\) In an article that describes advances made in assembling large gene constructs since the *M. mycoides* project, Gibson explained that assembling large gene fragments in host cells like yeast and *E. coli* is challenging because scientists cannot control what occurs inside of those cells.\(^{24}\) First, the cells can only accommodate so much DNA. *E. coli*, for instance, can only accommodate around 300 kb, while yeast cells can accommodate around 2 Mb.\(^{25}\) Second, some gene sequences can be toxic to host cells, which can prevent host cells from assembling the large gene fragments altogether.\(^{26}\) For example, genomes that are high in adenine and thymine, such as the *M. mycoides* genome, tend to be toxic in *E. coli*. According to Glass, whereas *E. coli* tolerated fairly large pieces of the *M. genitalium* genome, it did not tolerate pieces of *M. mycoides* genome that were the same length, perhaps because of the higher AT content of the latter

\(^{21}\) Gibson, e-mail message to author.
\(^{22}\) Glass, interview.
\(^{23}\) Gibson et al., “Creation of a Bacterial Cell.”
\(^{24}\) “Synthetic Life: Dr. Dan Gibson Discusses” (see chap. 4, n. 21).
\(^{25}\) Ibid.
\(^{26}\) Ibid.
genome. To overcome this problem, Gibson used yeast to recombine the 1,080 bp cassettes into 10,000 bp DNA assemblies, not *E. coli*. Third, *E. coli* has been known to recombine or mutate the synthetic fragment.

While Gibson was successful in assembling the *M. mycoides* genome fragments in yeast, my interviews show that one of the most difficult aspects of the assembly stage of the project occurred when the researchers extracted large pieces of the *M. mycoides* genome from yeast. Because large pieces of DNA are more brittle and difficult to handle than small pieces, Dr. Chuang and his team had to develop a method to extract the 100,000 bp pieces of *M. mycoides* DNA from yeast while also ensuring that they remain intact. The team’s first course of action was to consult the literature to identify the best pH for purifying DNA from yeast. This led them to the protocol for Qiagen’s Large-Construct Kit, a commercially available product used to purify large constructs of DNA. Using this protocol as a starting point, the team experimented with different pH levels to find one that would denature the yeast DNA while also leaving the *M. mycoides* DNA intact, a tedious process that involved repeated experimental cycles. Chuang explained,

> In yeast . . . you have to develop conditions so you don’t kill or denature your circular DNA, and you have to fine-tune the conditions, so this is

---

27 Glass, interview.
28 Gibson, “Assembling and Engineering.”
29 Ibid.
30 Glass, interview.
kind of the difficult part. So we tried different pH to denature the yeast chromosome but still keep the circular DNA that we’re interested in intact. So to find those conditions is very difficult. . . . We pretty much just followed this protocol and just make [changes] systematically based on pH to denature yeast chromosomal DNA. . . . So, pretty much, we just played around with the pH condition and tried to find the optimal pH that you can recover higher circular DNA with a lower contamination of the yeast chromosome.32

Although researchers were able to achieve an acceptable pH by using this process, Chuang explained that the purity of the resulting *M. mycoides* DNA was still low. He shared, “Even with purified DNA, when you run it on a gel, you can still see the contamination of yeast chromosomal DNA. That’s our speculation that, even though we can assemble the eleven pieces of the [*M. mycoides* genome] to create circular complete genome, the efficiency actually is . . . my recollection is probably less than 2 percent . . . that’s extremely low. Of course, all we need is just one [copy of the complete genome], but we believe one of the reasons is that the quality of the DNA in purified form is still not good enough.”33 Thus, although the team was eventually able to purify *M. mycoides* DNA from yeast, the results were not optimal. Glass noted, “There was a period between 2010, when we synthesized the whole [*M. mycoides* genome] and early 2012, when we couldn’t reproduce it, because we were unable to purify these pieces from

---
33 Ibid.
yeast. And we think that some of the kits we were using had been changed just a little bit. We were at such a knife edge of it working and not working.”

According to Glass, the researchers’ attempts to replicate this component of the *M. mycoides* genome assembly were confounded by changes in the reagent kits they were using; specifically, the reagents in Qiagen’s Large-Construct Kit. This highlights how reagents are extremely important in producing consistent results, because even small differences might prevent researchers from carrying out proven or well-established experimental protocols. JCVI researchers have since resolved their difficulties purifying *M. mycoides* from yeast; however, they indicated that this new protocol only works for *M. mycoides* and is not effective for purifying other organisms from yeast.

**Transplanting the *M. mycoides* Genome**

In addition to experiencing problems during the assembly stage, my interviews reveal that JCVI researchers who participated in transplanting the *M. mycoides* into *M. capricolum* experienced great difficulty. The groundwork for transplantation had begun years earlier, when Dr. Carole Lartigue, a postdoctoral fellow who had done her doctoral work developing molecular tools to study several *Mycoplasma* species, joined the team. Lartigue focused on finding ways to transplant *M. mycoides* into *M. capricolum* because there were better tools for working with those organisms. However, Lartigue’s efforts to transplant *M. mycoides* into *M. capricolum* took two years, and she suffered many

---

34 Glass, interview.  
35 Ibid.  
36 Ibid.  
38 Glass, interview.
failures before she succeeded. Glass, Lartigue’s supervisor, said, “After two years of just seven days a week [of continuous work], she came into my office . . . saying she wanted to . . . work on a new project; she couldn’t do this anymore. The next day, everything changed. We got that first colony [indicating a successful transplant], and from there, we have improved this method greatly, so instead of getting one colony per experiment, we now get thousands.”^{39}

The results of the transplantation were subsequently published in the journal *Science* in 2007.^{40} When asked how Lartigue was able to get the *M. mycoides* to transplant into *M. capricolum*, Glass explained, “We tried lots and lots of different approaches. And we had suspicions of something we thought might work and [Lartigue] was just exploring condition space, but these were hard experiments to do with a lot of reagent prep for every experiment. . . . Everything you could possibly think of that might allow you to move a really big piece of DNA into a cell [we tried].”^{41}

Some of the problems Lartigue experienced occurred because *M. mycoides* and *M. capricolum* shared a common restriction system.^{42} While a *M. mycoides* genome grown in a *M. mycoides* cell would be methylated and therefore protected from *M. capricolum*’s restriction system, the synthetic *M. mycoides* genome that was grown in yeast lacked this protection.^{43} Lartigue and colleagues overcame this hurdle by

---

^{39} Ibid.
^{40} Lartigue et al., “Genome Transplantation in Bacteria.”
^{41} Glass, interview.
^{42} Bacteria use restriction systems to protect themselves from foreign DNA. Foreign DNA is recognized because it is not methylated (i.e., does not contain methyl groups), whereas the host cell’s DNA is methylated.
^{43} Gibson et al., “Creation of a Bacterial Cell.”
methylating the synthetic _M. mycoides_ DNA or disrupting _M. capricolum_'s restriction system. In a 2009 interview with *Nature Biotechnology*, J. Craig Venter noted that the methylation problem “took 20 people two years to solve.”\(^4^4\) The methods that produced the successful transplantation were immediately codified and used for similar experiments the lab was conducting in parallel. Glass said, “We were pursuing half a dozen different approaches at the same time and also doing [Mycoplasma] genitalium at the same time using even different methods, and the moment we got that first success, everything else dropped and all the effort went into improving that method.”\(^4^5\)

JCVI researchers also encountered difficulties during the transplantation stage when handling the large, brittle DNA fragments that were part of the experiment. The researchers had to develop new methods to handle the brittle DNA because no techniques existed.\(^4^6\) One of the methods the team adopted to was to pipette gently and use pipette tips with wide openings to allow the DNA to pass through unobstructed. Chuang explained, “Pipetting is always a problem where you can introduce physical shearing. In general, whenever the DNA molecule is big . . . —say, maybe greater than, say, 40 or 50 kb—when you pipette it, you potentially introduce damage. So, for example, when we purified those 100 kb [fragments] from yeast, we had to use a tip with a wide open tip, and every time we had to pipette very gently. This all depends on the size of DNA, . . . but overall, you have to be very careful when pipetting and handling DNA.”\(^4^7\)

\(^4^5\) Glass, interview.
\(^4^6\) Science Channel, "Creating Synthetic Life" (see chap. 3, n. 23).
\(^4^7\) Chuang, interview, September 19, 2013.
Although JCVI researchers placed great emphasis on the ability to pipette gently, Glass indicated that the task is primarily a matter of muscle memory, and it varies even among experienced researchers. As he explained,

Our genome transplanters are really good at this [keeping supercoiled DNA intact]... I sat in the same hood... with Carole [Lartigue], and we used the same reagents, and the only thing different was each of us had our own pipettes and plates. And I did a transplant in parallel with her, and she got, you know, 2,000 colonies [successful transplants] and I got 20— and I thought I was doing exactly what she was doing in pipetting slowly. And we know that big pieces of DNA...—when you start getting to half a million bp or 300,000 base pairs—they become very fragile, and the bigger it is, the more fragile it is. And we have developed these tricks... which, even [with] me doing these tricks, it’s still very much a magic hand sort of thing... And what we find is that, when new people come in to do this, it takes them a few weeks to get good at it.48

Glass’s inability to produce as many colonies as Lartigue demonstrates the importance of expertise and know-how, gained over years of hands-on work, in achieving successful results. It also reveals how ambiguous words—such as “gently” or “carefully”—that are sometimes used in verbal or written protocols have different connotations to different researchers, even those who are highly experienced, and may influencer researchers’ abilities to successfully reproduce an experiment.

48 Glass, interview
When asked how researchers become proficient in pipetting these large DNA fragments, Glass answered, “It’s speed of doing things [and] realizing how you manage your workspace. And it may come down to getting relaxed enough to very slowly and smoothly pipette these things, not shaking too hard. . . . And . . . we have done experiments to find out what are the critical factors in genome transplantation, and we have identified there are temperature factors that are really critical in terms of keeping the reagents at the right temperature.” Glass’s response shows how the ability to pipette in the precisely required manner is acquired only through hands-on practice; furthermore, Glass notes that other factors, such as the temperature of reagents, are also critical to successful transplantation. Elaborating on the importance of reagent temperature during transplantation, Glass said that successful transplantation relied heavily on the temperature in which recipient cells were grown. The researchers started out growing the recipient cells at 37 degrees Celsius but learned that growing them at 30 degrees Celsius increased transplantation success. This discovery “took some inadvertent leaving of incubator doors open to figure out” and “was an accident.”

Glass also noted that, while a protocol exists for transplanting *M. mycoides* into *M. capricolum*, JCVI researchers have not been able to apply the protocol to other organisms, even though they have been trying to do so for several years. Furthermore, Chuang indicated that although JCVI researchers have been performing transplantation for five years, they still have problems reproducing the efficiency of the transplantation.

---

49 Ibid.
50 Ibid.
51 Glass, interview.
process. According to Chuang, the transplantation process is highly unpredictable: even though the process is always performed by the same researchers, it may be successful one day—producing a lot of colonies that are indicative of successful transplantation—but unsuccessful the next day—unable to produce as many colonies—even though the researchers are following the same protocol both days. Perhaps more puzzling, JCVI researchers had at one point increased transplantation efficiency, but a few months later the efficiency level dropped again for no apparent reason. Chuang said, “They [the scientists doing transplantation] just get different results. . . . We really don’t know why.” Although Chuang and colleagues are not sure why the efficiency of transplantation is inconsistent, he explained that the transplantation process is lengthy, and each step is critical. As such, any difference encountered at any step of the process can affect the outcome. For example, one theory to explain the inconsistencies is that the bottles where the cells are grown are not being washed well enough, leading to contamination. According to Chuang, researchers must pay attention to every detail—even seemingly small ones like the cleanliness of bottles—during every step of the transplantation process.

Using PCR

My interviews with JCVI researchers also show that they had difficulties using PCR, a common laboratory technique—in widespread use for 30 years—that amplifies small amounts of DNA. JCVI researches used PCR after the M. mycoides genome was

52 Chuang, interview by author, February 12, 2005.
53 Chuang interview, February 12, 2005.
54 Ibid.
55 Ibid.
assembled to ensure that all eleven junctions (that is, where the eleven 100,000 bp fragments joined) were in the correct order. Though none of the JCVI researchers I interviewed recalled experiencing problems with PCR during either the *M. genitalium* or *M. mycoides* projects, they do experience regular problems with PCR in other experiments, demonstrating that it is not a simple process.

For example, Chuang described a recent, ongoing attempt to amplify a 1 kb gene that had an extremely high AT content, which can be difficult to amplify.\textsuperscript{56} Chuang first conducted a literature search to find other protocols for amplifying AT-rich sequences. The literature indicated that temperature is important for extending and annealing these sequences. Using this information, Chuang’s team modified the temperature of the PCR reaction. However, this modification only slightly improved the yield. The team’s next course of action was a costly one: to test different commercial PCR kits until they found one that was able to amplify the sequence. Chuang explained, “I don’t know what else we should try. I think the condition of PCR is important, and different companies have different proprietary recipes that might be critical. So far, we haven’t had good luck yet. I think we’ve tried four companies so far, but we have to keep trying.”\textsuperscript{57} This shows that, contrary to being foolproof or automated, PCR protocols are insufficient for amplifying all genes. This also shows the amount of variation among commercial PCR kits, if one might amplify a problematic sequence while others cannot.

\textsuperscript{56} Because the *M. mycoides* genome has a high AT content, it is possible this difficulty might be encountered with *M. mycoides*.

\textsuperscript{57} Chuang interview, September 19, 2013.
Another JCVI researcher reported inconsistencies with PCR when using thermocyclers. The researcher described a project that involved amplifying a large volume of DNA using an enzyme that a client had required be used for the project. The client had provided JCVI researchers with a protocol for the PCR reaction, and the enzyme was part of that reaction; the JCVI had a contractual obligation to precisely follow this protocol. According to the researcher, for some inexplicable reason, the enzyme produced a better yield when the PCR was run in old thermocyclers, rather than new ones. Since the old thermocyclers were slower than the newer ones, the experiment took much longer than it would have if new thermocyclers had been used.

While it is not known whether either of these problems with PCR played a role in the M. mycoides project, the interviews confirm Jordan and Lynch’s previous findings that PCR is difficult to use, despite the abundance of supposedly automated kits meant to streamline the technique and despite PCR’s ubiquitous use in the life sciences over the past 30 years. The interviews also indicate that the effectiveness of PCR protocols varies depending on the sequence that is being amplified, rendering the automated kits insufficient for use on all sequences. Furthermore, new PCR thermocyclers may perform less efficiently than older ones in some experiments, which further supports the idea that progress in technology does not equate to better results.

Thermocyclers are instruments that are used to amplify DNA by PCR. Samples with reagents are placed in the thermocycler, and the thermocycler automatically raises and lowers its temperature in a series of preprogrammed steps.

Anonymous JCVI researcher, interview.
Verifying Gene Sequences

My interviews indicate that the limitations of DNA sequencing technologies also affected the *M. mycoides* project. When the completed 1.08 million base pair *M. mycoides* genome was first transplanted into *M. capricolum*, the *M. capricolum* cells were not transformed into *M. mycoides* as expected, prompting the researchers to reevaluate their protocol.\textsuperscript{60} To troubleshoot, they constructed a semisynthetic cell. This involved constructing *M. mycoides* from wild type genes—naturally derived nonsynthetic or engineered genes—and then taking the wild type pieces out one at a time and adding synthetic pieces in their place.\textsuperscript{61} The method allowed researchers to identify which synthetic gene contained the error. Using this method, they determined that one of the genes that the Blue Heron had provided them—the *dnaA* gene, which is essential to DNA replication—contained an error. Glass blamed this error on the ambiguous results that DNA sequencing machines provide, which gene synthesis companies use to verify that the gene they are synthesizing matches the sequence requested by their customer. Glass explained,

Companies that sequence all the time, they don’t keep their Sanger sequencing reagents that long. [Gene synthesis company] Blue Heron didn’t do that. And so the C dock gets old, and as it gets old you start getting shoulders on the peaks where the Cs [cytosine, one the four main bases found in DNA] are next to another base. So the technology that reads the sequence files didn’t pick up that there was a deletion. So when

\textsuperscript{60} Gibson et al., “Creation of a Bacterial Cell.”
\textsuperscript{61} Gibson, “Assembling and Engineering.”
we went back and resequenced that piece, we saw right away that in $dnaA$, in the first gene in the genome, there was what we believed to be a lethal deletion mutant right there. And this took us months to figure out.\textsuperscript{62}

According to Glass, Blue Heron might not have changed the reagents in their DNA sequencer, which produced ambiguous results that the company did not catch before it shipped the JCVI’s order. It was not until the JCVI had problems with transplantation—and then undertook a time-consuming troubleshooting process—that the error was discovered. This illustrates that nearly a decade after commercial gene synthesis companies entered the market, and more than 20 years after the automated DNA sequencer was introduced, sequencers still produce reports that are too ambiguous even for experts to interpret, a finding that was also reported in chapter 4. These errors can have a significant impact on customers, requiring them to undertake troubleshooting experiments that can take months. It also demonstrates the role that reagents play in experimental success; in this case, old reagents produced ambiguous sequencing results, which produced a cascade effect of experimental failure and costly, time-consuming troubleshooting.

**Resolving Contingencies in the $M.\ mycoides$ Experiment**

My interviews with JCVI researchers indicate that they relied heavily upon tacit knowledge throughout the course of the $M.\ mycoides$ project. First, they relied on specialized skills acquired during previous related projects to carry out the $M.\ mycoides$ project. For example, the $M.\ genitalium$ genome assembly, which required the JCVI team

\textsuperscript{62} Glass, interview.
to develop novel protocols for constructing large DNA fragments, provided Gibson with skills and knowledge he later used to build the *M. mycoides* genome. By drawing on those skills, Gibson was able to assemble the *M. mycoides* genome “pretty quickly.”

Likewise, Lartigue’s prior experience working with various *Mycoplasma* species was invaluable to the project because it provided a protocol for transplanting a synthetic genome into a recipient cell. In addition, Lartigue’s experience in transferring genomes into recipient cells likely contributed to the success of the project because the transfer required specialized skills that could only be acquired through practice.

Second, the researchers drew upon the tacit knowledge of their more experienced colleagues to resolve difficulties. One scientist who worked with Gibson during the assembly of the *M. genitalium* genome and, in 2011, worked on other JCVI synthesis and assembly projects referred extremely difficult synthesis problems to Gibson. As the scientist explained, “We just email Dan and he can answer it.” This indicates that, having developed the protocols for assembling both the *M. genitalium* and *M. mycoides* genomes and having assembled each genome, Gibson gained a significant amount of tacit knowledge that he was able to use later help coworkers resolve difficulties in their projects. This finding is consistent with studies of the biotechnology industry which show that access to the scientists who develop a particular technology is crucial for individuals who wish to create products from that technology because they need access to the original developer’s know-how.

---

63 Gibson, e-mail message to author.
64 Anonymous JCVI researcher, interview.
65 Zucker, Darby, and Brewer, “Intellectual Human Capital” (see chap. 3, n. 86).
In addition to relying on tacit knowledge, JCVI researchers also consulted the literature and engaged in trial and error to resolve difficulties. Although their first course of action for resolving contingencies typically involved consulting the literature or other published protocols, JCVI researchers—like the gene synthesis company employees I interviewed—often had to engage in trial and error to adapt these protocols to their specific needs. This was the case, for example, when researchers in Chuang’s lab attempted to determine which pH was best for purifying *M. mycoides* DNA from yeast or which commercial PCR kit was capable of amplifying a problematic, AT-rich gene. As Glass observed, the team members often approached the trial and error process in different ways:

I noticed that Dan Gibson and Carole Lartigue, their approach is just brutal: get in and explore condition space. If they think something is almost working, . . . if they believe they are on their way to solving a problem, they will just get in every day and do it again and again and again. Slight variations, taking meticulous notes, until they get what they want. And it is painstaking and it is very effective. Ham Smith is like that, but he will try a much broader variety of things, but it is still this “consider all the possibilities and test the most likely ones first [approach].”

Some of this reliance on trial and error was likely due to the *M. mycoides* experiment’s novelty, which required them to develop their own protocols as they learned how to conduct the experiment, including a final protocol for transplanting the synthetic genome.

---

66 Glass, interview.
into *M. capricolum*, a protocol for purifying large DNA constructs from yeast chromosomal DNA, and a protocol for handling large DNA constructs.

In addition to tacit knowledge, trial and error, and explicit forms of knowledge, JCVI researchers also modified their laboratory technologies to resolve difficulties. For example, they purchased pipette tips with wide openings to prevent them from shearing large DNA fragments and used different commercial PCR kits to amplify a particularly problematic, AT-rich gene. They also used an older thermocycler to increase the efficiency of a particular PCR amplification. Table 5 summarizes the contingencies encountered during the *M. mycoides* project.

<table>
<thead>
<tr>
<th>Contingency</th>
<th>Example</th>
<th>Contingency resolution method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assembling large gene</td>
<td>Propagating sequences of certain lengths in yeast; toxicity in <em>E. coli</em> (that is, problems with host cells)</td>
<td>• Trial and error</td>
</tr>
<tr>
<td>fragments</td>
<td>Purifying <em>M. mycoides</em> genome from yeast cells</td>
<td>• Transferring knowledge (consulting literature)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Trial and error</td>
</tr>
<tr>
<td>Transplanting genomes</td>
<td>Handling large fragments of DNA (that is, the <em>M. mycoides</em> genome)</td>
<td>• Transferring knowledge (training via observation and replicating routines)</td>
</tr>
<tr>
<td></td>
<td>Determining the best temperature for incubating recipient cells</td>
<td>• Transferring knowledge (consulting literature)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Trial and error</td>
</tr>
<tr>
<td></td>
<td>Increasing transplant efficiency</td>
<td>• Increasing attention to detail during all aspects of the transplantation process</td>
</tr>
<tr>
<td>Using PCR</td>
<td>Amplifying AT-rich sequence</td>
<td>• Transferring knowledge (consulting literature)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Trial and error</td>
</tr>
<tr>
<td>Verifying sequences</td>
<td>Obtaining incorrect sequences from commercial supplier</td>
<td>• Trial and error (troubleshooting experiments)</td>
</tr>
</tbody>
</table>
What the Case Study Tells Us

This case study indicates that the *M. mycoides* experiment was difficult and that major components of the experiment—purifying large genes from yeast, transplanting genomes into recipient cells, pipetting large fragments of DNA, and obtaining correct sequences—are still problematic and require researchers with specialized knowledge and significant tacit knowledge. Moreover, certain aspects of the experiment, such as transplantation, are difficult for researchers to replicate consistently. The fact that JCVI researchers—the experts who developed the protocol—cannot apply it to other organisms indicates that malefactors would not be able to use the published protocol to create a select agent.

My interviews show that tacit knowledge greatly influenced the success of the experiment and that teams of individuals with unique expertise and skills are required to solve problems. Several JCVI researchers had been working with *M. genitalium* and *M. mycoides* or using similar experimental methods for more than fifteen years. This experience gave the researchers unique insight into the nature and behavior of the organisms and may have helped them resolve problems more quickly. The assembly of the *M. genitalium* genome, which occurred two years before that of *M. mycoides*, conferred Gibson with specialized knowledge that allowed him to more effectively assemble the latter’s much larger 1.08 Mb genome. In fact, Gibson’s colleagues still rely on his knowledge of assembly to resolve their difficulties, even though some of these colleagues also have experience assembling large genes. Likewise, Lartigue’s extensive prior work on *Mycoplasma* species contributes to the success of genome transplantation,
conferring her with knowledge about *Mycoplasma* species and the tacit skills required to transplant the *M. mycoides* genome successfully. These latter skills have proved difficult to acquire, as my interviews with Chuang interviews attest: five years later, JCVI researchers still encounter difficulties transplanting the *M. mycoides* efficiently.

These examples show how—nearly five years after the experimental protocol was published and optimized—JCVI’s highly experienced researchers still encounter difficulties that can only be resolved by colleagues with more expertise and tacit knowledge about specific methods. Moreover, the skills that researchers on the assembly and transplantation teams possess are not easily transferred to other aspects of the project. In other words, Gibson would not be able to perform genome transplantation as effectively as Lartigue, and Lartigue would not be able to assemble a large genome as effectively as Gibson. No one person on the JCVI team has all of the skills required to successfully replicate the experiment. This is consistent with Ben Ouagrham-Gormley’s findings that knowledge reservoirs are interdependent and that access to only one reservoir limits an individual’s or organization’s ability to replicate previous work.67

In addition to influencing researchers’ understanding of the organisms they were working with and the techniques they were using, tacit knowledge may have also influenced the way these researchers used laboratory instruments. For example, when Lartigue and her supervisor conducted transplantations side-by-side, Lartigue generated a vastly greater number of successful transplants, even though her supervisor took great care to follow Lartigue’s movements and pipette carefully. This indicates that her

---

67 Ben Ouagrham-Gormley *Barriers to Bioweapons*, 27.
supervisor lacked the tacit skills—gained by performing the technique repeatedly—to pipette as gently as Lartigue, who had much more practice with that particular task. Lartigue, for her part, might also have lacked the ability to demonstrate or explain what constitutes gentle pipetting. This finding is consistent with previous literature on the tacit knowledge involved in successfully using or maintaining laboratory instrumentation and equipment: researchers who are perceived as having a knack for using a specific instrument or carrying out a protocol are often described as having “lab hands.”

The concept of “lab hands” was echoed by Glass when he explained that, although researchers had “tricks” to help them pipette the genome, the transplantation process is “still very much a magic hand sort of thing.”

Another important lesson is that the errors contained in synthetic genes obtained from commercial companies can have a devastating effect on the users, leading to experimental delays and failures, and researchers must have the skills, expertise, and material resources to identify those errors. In JCVI’s case, several months of additional work were necessary to determine that one of the genes obtained from Blue Heron contained a deletion in an essential gene, causing a cascade of additional errors.

The methods that JCVI scientists use to overcome contingencies are similar to those used by gene synthesis firms. However, gene synthesis firms prefer to synthesize shorter and predictable sequences so they can avoid trial and error, whereas the JCVI

---

68 Doing, “‘Lab Hands’ and the ‘Scarlet O,’” (see chap. 4, n. 74).
69 Glass, interview.
relied heavily on trial and error to assemble and transplant the *M. mycoides* genome as well as to overcome other difficulties commonly encountered in their laboratory.

**Implications for National Security**

Even though the scientists who worked on the *M. mycoides* project are experts who possess decades of experience working with similar organisms and techniques, the project was fraught with much difficulty. Some aspects of the experiment—namely, the purification of large constructs of DNA and the transplantation of large DNA constructs into host cells—continue to be a problem for JCVI researchers well after the publication of the 2010 paper, indicating that methods for resolving errors are not emerging as quickly as some national security analysts suppose. In addition, no one person on the JCVI team has all of the skills required to replicate the experiment. Moreover, the experiment is specific to *M. mycoides*, and the JCVI researchers who developed the method have not been able to adapt it for use with other organisms thus far. Additionally, assembling a large genome is cost prohibitive; in 2014, it cost more than $1 million to build a 1 Mb bacterial genome like *M. mycoides*.  

This case study shows that any individuals or groups who intend to create a synthetic pathogen from DNA purchased from a gene synthesis firm would need to rely on a more applicable protocol—assuming one exists—or conduct lengthy trial-and-error experiments in a stable research environment to develop their own novel method, which implies that they also possess the relevant expertise and skills. Each stage of the development of such protocols is fraught with uncertainty, even for experts in the field. If

---

70 “Synthetic Life: Dr. Gibson Discusses.”
malefactors wish to pursue this course, they would first need to synthesize the genes themselves or order them from synthesis companies. Because synthesis companies routinely screen orders to determine if they match sequences in select agents, the most likely option would be for malefactors to synthesize the genes themselves. Interviews with gene synthesis firms (see chapter 4)—experts in the field—show that even they have difficulty synthesizing particular sequences. My interviews with the gene synthesis firms and the JCVI researchers both show that verifying that sequences are correct is a difficult process, and receiving incorrect sequences can lead to significant experimental delays and failures.

Assuming malefactors overcome these hurdles, they would then need to assemble the organism’s genome, a difficult undertaking that would increase in complexity if the genome were large. The sequence of the resulting genome would then need to be verified. If this was successful, the malefactors would then need to transplant the genome into a cell. Because protocols for genome transplantations are specific to each organism and genome, malefactors would need to conduct extensive trial and error to develop and test a protocol that would work for their pathogen of choice. Assuming they could develop such a protocol, implementing it during the transplantation stage would be difficult if the genome were large, due to the fragility of large DNA fragments.

As an interviewee from a gene synthesis company pointed out, if a malefactor wished to create a select agent or other pathogen, he or she could do so much more quickly using standard genetic engineering and wild type genes. Commenting on the *M. genitalium* study, the interviewee said,
The stuff that the Venter team did is pretty cool; it shows that you can make huge chunks of DNA. From a manufacturing standpoint, it’s pretty cool, but from a science standpoint it’s fairly mundane . . . because he could have done the same thing in two weeks by taking wild type *M. genitalium* and making some mutations on it. He picked the most complicated possible way to make the genome, which is fine, but where’s the value? How can you go to Pfizer with the genome and ask them for money for something that was made in a very complicated manner, which could have been made in a weekend?  

Thus, although the JCVI work may have appeared to make it easier to create a pathogen synthetically, there are other methods for doing so that are less expensive and time-consuming, but they may be just as difficult to carry out from a technical standpoint. This was demonstrated by a recent analysis of past state and terrorist bioweapons programs; security researcher Ben Ouagrham-Gormley has shown that using classical agents is also very difficult.  

Therefore, many national security analysts’ assumption that the threat from bioterrorism is heightened because of advances made in synthetic biology, such as those that occurred during the *M. mycoides* project, dramatically underestimates the difficulties encountered in such projects and ignores the importance of having access to the appropriate skills and expertise.  

---

72 Ben Ouagrham-Gormley, *Barriers to Bioweapons*. 

180
CHAPTER 6
RESEARCH FINDINGS AND POLICY IMPLICATIONS

My interviews affirm the findings of previous STS research on contingencies in the sciences, as well as the ways that scientists acquire and transfer tacit knowledge. Specifically, the interviews show that (1) synthesizing genomes is technically more complex than threat assessments make it out to be; (2) published protocols provide insufficient information, making them difficult to use even for experienced scientists; (3) researchers rely heavily on their tacit knowledge to overcome difficulties; and (4) most of the current threat assessments ignore these factors and overstate the threat that malefactors could use synthetic genomics to create pathogens. These observations apply to highly automated, industrial technologies like synthetic genomics, as well as novel scientific research like synthetic biology. Even with their vast resources, gene synthesis firms still encounter difficulties. Furthermore, research organizations like the JCVI, which specialize in conducting novel research, have to engage in long and painstaking trial and error throughout the experimental process to resolve difficulties. Together these findings show that terrorists or other malefactors who might attempt to use synthetic genomics and synthetic biology to develop a biological agent will encounter significant challenges throughout the process. In this chapter, I discuss these four findings in greater detail.
Finding 1: Synthesizing genomes is a technically complex undertaking that terrorists are unlikely to master

As mentioned previously, research has shown that gene synthesis can be too difficult for scientists to carry out, so they often outsource synthesis to commercial firms.\(^1\) Although mainstream threat assessments are correct that the level of automation in commercial gene synthesis is high, my interviews indicate that approximately 10 percent of synthesis orders present some type of technical difficulty and that key stages of commercial synthesis—such as assembling large genes, cloning genes into vectors, and interpreting chromatograms from DNA sequencers—require human skills and sometimes unique knowledge to ensure successful results. Many of the difficulties encountered during gene synthesis are due to the unpredictable nature of DNA itself. The synthesis firms I interviewed relied on the tacit knowledge and skills of their more experienced employees to overcome many of these difficulties. In addition, because scientists cannot predict the behavior of all DNA sequences, synthesis firms’ design software cannot identify and design around problematic sequences. Therefore, terrorists who wish to synthesize genes would not only have difficulties predicting the behavior of DNA, but they also would not be able to use automated systems to solve complex problems that are beyond the systems’ capabilities. As a result, unless the terrorists have the specialized skills and expertise required, the likelihood that they will successfully synthesize pathogen genes is slim.

My interviews with the JCVI researchers who participated in the *M. mycoides* experiment show that it is difficult to assemble and then purify large genomes from yeast

---

\(^1\) Marsic et al., “PCR-Based Gene Synthesis” (see chap.1, n. 14).
and that handling those genomes, such as during the transplantation stage, requires great care. Gibson, the JCVI researcher who assembled the *M. mycoides* genome, had acquired knowledge about how to assemble large genomes during the previous *M. genitalium* experiment, and that knowledge enabled him to assemble the *M. mycoides* genome more quickly. Gibson has since codified this knowledge by publishing his assembly method and also by partnering with SGI-DNA, Inc. to launch a line of reagent kits based on his Gibson Assembly™ method. The Gibson Assembly™ method is based on a 2009 protocol that Gibson and colleagues published in *Nature Methods* and involves what DNA synthesis company Integrated DNA Technologies describes as “directional cloning of multiple DNA fragments in a single reaction,” allowing researchers to assemble large genes “rapidly.”

According to an article on the Integrated DNA Technologies website, the Gibson Assembly™ method “takes only one hour,” and the upper size limit for assembled sequences is “>1Mb.” SGI-DNA, Inc., which manufactures the reagent kits, claims the method has been cited in over 600 scientific publications, suggesting that it is scientists’ method of choice for assembling large constructs. According to one reagent supplier, the synthetic biology community rapidly adopted the method because of its “ease-of-use, flexibility and suitability for large DNA constructs.” Indeed, Gibson affirmed that the methods and technologies for assembling large DNA fragments are

---

2 Gibson et al., “Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases,” *Nature Methods* 6, no. 5 (May 2009): 343–5, doi: doi: 10.1038/nmeth; and “Synthetic Life: Dr. Dan Gibson Discusses” (see chap. 4, n. 21).
3 “Synthetic Life: Dr. Dan Gibson Discusses.”
continuing to improve; he asserted that assembling *M. mycoides* would take weeks using current methods, whereas it took months to achieve in 2010.”

However, these claims about the Gibson Assembly™ method are similar to claims about PCR in the late 1990s. There is evidence that the Gibson Assembly™ method, like PCR, is not as easy to use as the manufacturer websites would have their customers believe. A subpage on the widely used online public forum Reddit® that is dedicated to helping scientists crowdsource solutions to their laboratory problems shows a number of questions about how to use the method, an indication that it is not easy to use. One user of the forum, whose screen name is Ch1b0, submitted the following question:

So here is the problem. I have been trying to get a Gibson Assembly reaction to work for what seems like an eternity now. So far all of my transformed cells only contain my vector without my insert(s). The gene I am trying to assemble into my vector is only 2.2kb in length and my vector is ~5kb linearized. Unfortunately for me, I have multiple bands when amplifying my gene of interest, so I am forced to gel extract and purify that PCR product. As I have never really gotten a gel extraction to work effectively and only get a yield of 10-20ng/ul. I have also tried to

---

6 Gibson, e-mail message to author (see chap. 5, n. 4).
7 Jordan and Lynch, “The Dissemination, Standardization, and Routinization” (see chap. 1, n. 11).
8 Reddit is a website where registered members can submit content, making it an online bulletin board system. The site hosts a public subpage (a “subreddit,” or community) called “labrats,” which serves as an open forum for laboratory scientists who want to share stories, vent their frustrations, or solicit help for their experiments. The subreddit can be accessed at the following URL: https://www.reddit.com/r/labrats/.
amplify my gene in two halves and assemble them in my vector [in other words, to split the gene up to make it easier to assemble]. Again, failure. I am at my whits [sic] end here and getting very frustrated. Have any of you every [sic] used this method of gene fusion? What pointers/advice can you give me?\textsuperscript{10}

The responses to this inquiry varied; most respondents provided advice on how to use the Gibson Assembly\textsuperscript{TM} method, while others commiserated about the difficulties associated with PCR gel product purification and offered recommendations for how to improve that process.\textsuperscript{11} The first commenter who responded to Ch1b0’s question offered several suggestions, then noted that he or she had worked with Gibson at the JCVI, which lent credibility and authority to the answers and showed that the commenter had hands-on experience with an expert.\textsuperscript{12} This exchange is shown in figure 10.

\begin{footnotesize}
\begin{enumerate}
\item Ch1b0 [Reddit User], “Help! I’ve Been Troubleshooting Gibson Assembly for Months!” \textit{Reddit, Labrats Community}, accessed November 6, 2015, https://www.reddit.com/r/labrats/comments/1x8pz6/help_ive_been_troubleshooting_gibson_assembly_for/.
\item Ibid.
\item MrTypie [Reddit User], comment on Ch1b0, “Help!”
\end{enumerate}
\end{footnotesize}
The questions posed by Ch1b0 and other Reddit users about how to use the Gibson Assembly™ method indicate that Gibson was not able to fully translate all of the tacit knowledge he acquired during the *M. mycoides* assembly into Gibson Assembly™ protocols or automated kits. This is likely because Gibson does not know what tacit knowledge he has about assembly that is essential for a successful outcome; therefore, he may have inadvertently omitted key details. This is consistent with Collins’ research, which asserts that tacit knowledge is transferred most effectively in person, when a knowledge recipient mimics the actions of a donor, because these interactions allow researchers to transfer different types of tacit knowledge, such as knowledge they unwittingly omitted from protocols (concealed knowledge), knowledge they do not know.

---

13 Ch1b0, “Help!”
they possess (unrecognized knowledge), and knowledge that is best conveyed by pointing and demonstrating (ostensive knowledge). Some of the difficulties with the Gibson Assembly™ method might be due to the relative newness of the technique and associated kits; Gibson and colleagues first published the method in in 2009, so the kits have been on the market for no more than five years. However, if PCR is any indication, users will continue to experience for several more years at least.

In addition to problems assembling and purifying large gene fragments, JCVI researchers also had problems handling the large \emph{M. mycoides} genome because it would break if they did not pipette it gently enough in the transplantation stage. They modified their laboratory tools to prevent the genome from breaking, but their ability to successfully transplant the genome ultimately relied on their ability to pipette gently, a skill that was acquired through practice and muscle memory. This is reminiscent of Vogel’s finding that using the glass Dounce homogenizer to break up HeLa cells during the poliovirus synthesis project was more of an art than a science; the cells needed to be broken up just enough, but not too much. The researchers who used the Dounce homogenizer successfully told Vogel that using it should feel or look a certain right way, vague phrases that are impossible to communicate in protocols.

The JCVI’s researchers have also encountered another contingency that they cannot explain: five years after they published the results of their \emph{M. mycoides} experiment, the efficiency of their genome transplantation procedure is still low, and they

\footnote{Collins, “Tacit Knowledge, Trust, and the Q of Sapphire” (see chap. 2, n. 146).}

\footnote{Vogel, “Framing Biosecurity” (see chap. 3, n. 35).}
are not sure why. At one point, the efficiency increased for no known reason but decreased a few months later, even though the researchers were using the same protocol when the efficiency was high as they were when the efficiency dropped back down. Again, this is similar to Vogel’s findings that scientists involved in the poliovirus synthesis had trouble producing and maintaining the HeLa cell-free cytoplasmic extracts, despite the fact that procedures had been available for more than 20 years. Even though the researchers always used the same procedures, they were still not able to replicate the results reliably. This shows that there are some factors to experimental success that even researchers cannot account for.

Perhaps one of the more surprising findings of my research is the fact that PCR is still difficult for researchers to use. Jordan and Lynch demonstrated that this was the case in 1998, when PCR had been in use for 15 years. Other researchers have noted continued difficulties in recent years. My interviews affirm these previous findings by showing that, nearly 30 years after its widespread adoption, PCR is still procedurally flexible; in other words, it varies depending on the sequence that is to be amplified, the equipment and reagents at hand, and other considerations. JCVI researchers had problems amplifying sequences with an AT-rich content, such as those in the M. mycoides genome, and found that some sequences and enzymes amplify better when their PCR reactions are carried out in old thermocyclers. To overcome the problem of amplifying genes with AT-

---

16 Vogel, Phantom Menace, 77.
18 Jordan and Lynch, “The Dissemination, Standardization, and Routinization.”
19 Vogel, Phantom Menace; and Ben Ouagrham-Gormley, Barriers to Bioweapons.
rich content, JCVI researchers had to test different manufacturers’ reagent kits until they found one that worked, a time-consuming and expensive process.

The same online resource for laboratory troubleshooting that demonstrated problems with the Gibson Assembly™ method also shows problems with PCR. Though the examples are numerous, one exchange is especially informative about PCR’s procedural flexibility. In it, user biomedthrowaway1 asks for help explaining possible causes of failure with PCR and how to go about testing these causes of failure. Another user responds as follows:

**a_karenina:** I think I could write an entire thesis on everything that can go wrong with PCR and potential troubleshooting . . . at least give us your thoughts first before we do your homework for you. You won’t learn anything that way. . . .

**biomedthrowaway1:** I’ve genuinely spent the last 2.5 hours reading through my notes and have gotten nowhere—my initial thoughts were problems related to primer design, but given that they’ve [the user’s professors] already given us the primers to work with, it can’t be that. They gave us the protocol, including the PCR temperatures [*sic*] and all of the reactants are already in their correct concentrations. Apart from human error, I just don’t see what could go wrong. . . .

**a_karenina:** Often, PCR is more of a magical art than a science. Just because you have all the right reagents and perfectly designed primers
doesn’t mean it will work the first time. There are many different aspect
[sic] that can go wrong.\textsuperscript{20}

Respondent a\_karenina went on to explain the various problems with primers and the
template (including whether it was GC-rich). Another respondent provided a list that
showed eleven things that could go wrong during PCR then noted, “This is not an
exhaustive list. There are many potential variables and therefore many potential issues
that can arise when conducting PCR. A reaction designed on the computer won’t always
work the first time and the parameters will have to be determined empirically.”\textsuperscript{21}

Not only does this thread show the many ways in which PCR protocols are
procedurally flexible, it also shows that inexperienced scientists may have misperceptions
about the effectiveness of protocols and reagents.\textsuperscript{22} Biomedthrowaway1 is likely a
student, because he or she was asked to identify possible causes of failure before
conducting an amplification later that day, and a\_karenina, apparently a more
experienced scientist, identified this exercise as “homework.” Tellingly,
biomedthrowaway1 was not able to come up with causes of failure because students were
given “the primers to work with . . . the protocol, including the PCR temperatures and all
of the reactants . . . in their correct concentrations.”\textsuperscript{23} Although biomedthrowaway1 did
mention that human error could potentially cause the failure, this was only after he or she
evaluated the suitability of the reactants and the availability of primers and a protocol.

\textsuperscript{20} Biomedthrowaway1 [Reddit User], “Possible Causes of PCR Failure?” \textit{Reddit, Labrats Community},
\textsuperscript{21} Ibid.
\textsuperscript{22} Lynch, “Protocols, Practices, and the Reproduction of Technique” (see chap. 1, n. 5).
\textsuperscript{23} Biomedthrowaway1, “Possible Causes?”
This indicates that, like many national security analysts, biomedthrowaway1 was unaware that a laboratory technique could fail when protocols and reagents are available. User a_karenina’s response that PCR is often “more of a magical art than a science”\textsuperscript{24} is consistent with input from one gene synthesis firm representative that some highly skilled people in science have “hands that don’t work right”\textsuperscript{25} and JVCI scientist Glass’s statement that genome transplantation is “still very much a magic hand sort of thing.”\textsuperscript{26}

Another interesting finding from my interview with JCVI researchers was that some individuals within the organization had expertise in specific aspects of the experiment, such that no one individual on the team possessed all of the expertise needed to carry out the entire experiment. For example, Gibson acquired expertise in assembling genomes, while Lartigue acquired expertise in transplanting genomes—two skill sets that are not directly related. Neither researcher would be able to carry out the other person’s research as effectively. My interviews with gene synthesis firm representatives indicate that some of the knowledge that is required to synthesize genes is also specialized. For example, some employees are better at designing genes or synthesizing difficult sequences than others. This indicates that terrorists who wish to synthesize genes or replicate the \textit{M. mycoides} experiment would need to develop expertise in multiple laboratory techniques, each of which might require a different set of skills and could be very difficult to master.

\textsuperscript{24} Ibid.
\textsuperscript{25} Company A, interview, April 5, 2011 (see chap. 4, n. 34).
\textsuperscript{26} Glass, interview (see chap. 5, n. 5)
As demonstrated by previous STS research discussing contingencies in the biological sciences, laboratory troubleshooting forums, and my interviews, life sciences research is rife with difficulties that are not reflected in either threat assessments of technologies or many company websites that attempt to market their automated technologies as foolproof. However, another perhaps more impactful source of information also fails to mention these contingencies: the protocols themselves.

**Finding 2: Published protocols are often not detailed enough for even experienced scientists to use**

As shown in chapter 3, threat assessments on biotechnologies often assume that terrorists or other malefactors could misuse the technologies so long as they have access to materials, protocols, and individuals with the “required skills.” However, previous STS research shows, and my interviews affirm, that it is often not a straightforward process, nor do researchers always achieve success, when they follow published or internal protocols and use associated laboratory tools. For example, gene synthesis firms have access to automated instruments and rely heavily on internal SOPs and database rules to fill most of their orders. However, approximately 10 percent of orders are difficult to synthesize due to the unpredictable nature of DNA and the sensitivities of cloning vectors. While most companies have developed SOPs to help them handle difficult sequences, these are not always sufficient, so the difficult sequences must be transferred to more experienced scientists within the company, such as individuals in R&D departments. The companies’ gene design software often does not identify these problems beforehand, and the software itself presents a problem because new knowledge
acquired from the literature or from resolving problems with difficult sequences must be programmed into the design database by a human expert. Moreover, the firms’ SOPs are not sufficiently detailed to teach new employees how to synthesize genes. Instead, employees must receive hands-on training with more experienced scientists before they can work independently, a process that can take as long as six months.

JCVI researchers reported using protocols to resolve difficulties in the laboratory, but these protocols were often modified through trial and error to suit the researchers’ needs. For example, the researchers consulted the literature to figure out how to amplify an AT-rich sequence using PCR but were still unable to amplify the problematic sequence despite following the literature’s guidance. As a result, the researchers resorted to testing different manufacturers’ PCR kits to find one that might work. In addition, some of the JCVI’s own protocols required significant interpretation. For example, protocols that emphasized the importance of gentle pipetting were not sufficient to allow researchers to achieve successful genome transplantation. Instead, the individual who had the most experience handling these large DNA fragments was more successful at transplanting the synthetic M. mycoides genome than her less experienced counterparts were because she had acquired the muscle memory to know what set of movements constituted “gentle” pipetting.

In addition to my interview findings, questions that appear on troubleshooting forums for life sciences research, shown above, reinforce the finding that protocols and materials such as automated kits are insufficient for allowing successful PCR or large-gene assembly. Indeed, and the very existence of these troubleshooting forums
demonstrates that these problems are widespread but taken-for-granted aspects of doing science.

Given this information, it is clear that access to materials and explicit knowledge are insufficient for achieving experimental success because they omit key pieces of the authors’ tacit knowledge, such as the knowledge they do not know they possess and knowledge that can only be understood physically via pointing, demonstrating, or feeling.\textsuperscript{27} Thus, the protocols do not include all of the knowledge that their authors possess, requiring significant interpretation from readers. These omissions therefore require that scientists have access to either a prior knowledge base that can help them carry out the protocol or, ideally, that they work closely with a study’s authors.\textsuperscript{28} As one biotechnology writer noted, “narrative descriptions in the methods sections of research articles or in laboratory notebooks don’t capture all of the individual manipulations of an experimental system or the metadata from instruments that could help to explain variations.”\textsuperscript{29} In addition to requiring significant interpretation, the STS research and my interviews show that researchers must adapt the information in these protocols to suit the specific needs of their experiments, which often involves tedious trial and error.\textsuperscript{30}

The omission of key knowledge in protocols and other explicit forms of knowledge might be at least partially to blame for the current irreproducibility crisis in the life and social sciences. In 2010, pharmaceutical giant Bayer Healthcare reported that its scientists could not reproduce approximately 75 percent of published research in

\textsuperscript{27} Collins, “Tacit Knowledge, Trust, and the Q of Sapphire.”
\textsuperscript{28} Ben Ouagrham-Gormley, \textit{Barriers to Bioweapons}, 11.
\textsuperscript{30} Ben Ouagrham-Gormley, \textit{Barriers to Bioweapons}, 27.
cancer, cardiovascular disease, and women’s health. A year later, researchers from pharmaceutical company Amgen and the M.D. Anderson Cancer Center reported that they could only replicate six out of 53 landmark studies in cancer biology—an 11 percent reproducibility rate. In 2012, scientist Glenn Begley, who led Amgen’s hematology and oncology group when the replication studies were carried out, coauthored a commentary in the journal *Nature* with a cancer researcher from the M.D. Anderson Cancer Center to discuss the reproducibility problems. According to the article, Amgen researchers visited the laboratories that had carried out the original studies and exchanged reagents with them when they could not reproduce findings, but the resulting reproducibility rates were still low. Authors of the few studies that could be reproduced had paid “close attention to controls, reagents, investigator bias and describing the complete data set.” Authors of studies that could not be reproduced typically presented the results of only one of their experiments or presented findings of experiments that supported their hypotheses but did not necessarily reflect their entire data set.

These alarming findings have spawned several efforts to examine or mitigate the irreproducibility problem. ScienceExchange, an online service that helps scientists outsource their research to other institutions, started the Reproducibility Initiative to help

---

31 Biotechnology and pharmaceutical companies often repeat experiments from studies that show promise for further development.
32 S. Webb, “Supporting More Reliable Results.”
34 Ibid., 531.
researchers validate their findings. In this initiative, ScienceExchange matches submitted experiments with an appropriate laboratory, which then attempts to reproduce the study’s findings. Researchers whose findings have been reproduced can boast that their studies have been “independently validated,” boosting their credibility. In 2013, the Center for Open Science was founded. According to its website, the Center supports research on scientific practices, fosters open science communities, and offers a free cloud-based service called the Open Science Framework to help researchers manage their research data. University of Virginia psychologist Brian Nosek is the executive director of the Center. In 2015, as part of the Open Science Framework’s Reproducibility Project: Psychology, Nosek and nearly 270 of his colleagues repeated the work of 98 papers from three psychology journals to see if they achieved the same results as the studies’ original authors. Only 39 replications were successful. The researchers communicated with studies’ original authors throughout the process to ensure there was no known difference between the original protocol and the reproduction protocol. According to Nosek, “There’s a lot of learning that happens when you write up the protocol and share it with the original team” because the original team invariably says critical details were left out, even though “none of that [critical information] was available in the original materials.” Nosek reports that original authors are “almost always” surprised to learn

36 The Reproducibility Project: Psychology aims to estimate reproduce reproducibility of psychological science.
38 Briak Nosek, interview by author, September 14, 2015.
that their studies cannot be reproduced, and reported only one experience where an
original author admitted to publishing only part of their results, indicating that
researchers do not intentionally omit critical data.\textsuperscript{39} Nosek also reported that the
Reproducibility Project: Cancer Biology initiative often heard about the importance of
having “lab hands,” reinforcing the notion that tacit knowledge in the form of muscle
memory is an important aspect of life sciences research.

In addition to the Open Science Framework, Riffyn, a laboratory software
company located in Oakland, California, is offering cloud software to help researchers
record and share data about their experimental protocol and integrate analytical data.\textsuperscript{40}
The purpose of the tool is to help make researchers’ tacit knowledge explicit and to share
that body of knowledge openly with other scientists.\textsuperscript{41} Another initiative to reveal tacit
knowledge in the sciences is the \textit{Journal of Visualized Experiments (JoVE)}, a peer-
reviewed scientific journal that videotapes certain experimental protocols. The \textit{JoVE} was
established in 2006 and, since then, has published more than 3,000 video articles in
numerous fields, including bioengineering, biology, chemistry, engineering,
neuroscience, immunology and infection, and medicine.\textsuperscript{42} By using a video format, \textit{JoVE}
gives authors the opportunity to more clearly present their methods. In providing clear
descriptions of methodologies, \textit{JoVE} hopes to help scientists overcome problems with

\textsuperscript{39} Ibid.
\textsuperscript{40} S. Webb, “Supporting More Reliable Results.”
\textsuperscript{41} Tim Gardner, interview by author, September 21, 2015.
\textsuperscript{42} “About Jove, A New Movement in Scientific Publishing,” \textit{Jove}, last modified 2015,
http://www.jove.com/about.
reproducibility and reduce the amount of time they spend learning new techniques.\textsuperscript{43} Users report that JoVE articles have helped clarify other researchers’ experimental methods.\textsuperscript{44}

Dr. Tim Gardner, chief executive officer of Riffyn, founded the company after years of frustration trying to reproduce experiments in biotechnology laboratories. According to Gardner, the problem with irreproducibility is the lack of transparency in scientific protocols. In an interview with O’Reilly Media’s \textit{Radar} blog, which focuses on emerging technologies, Gardner discussed the lack of detail in scientific papers by saying, “The problem is there’s an oral tradition in labs of how information is kept, transferred, and taught. It’s much more like apprenticeship in a blacksmith shop than like a precision engineering and training operation. . . . This problem is widespread.”\textsuperscript{45} This indicates that Gardner’s experience in the laboratory has been similar to the ideal tacit knowledge transfer method that Collins describes, in which a knowledge recipient and a knowledge donor engage in an apprenticeship relationship and, as the recipient physically replicates the donor’s routines, the donor transfers his or her ostensive knowledge by demonstration.\textsuperscript{46}

Discussing irreproducibility, Gardner said, “From a technical perspective, obstacles [to reproducibility] are the lack of blueprints of how experiments are

\begin{footnotesize}
\begin{footnotes}
\item[43] Ibid.
\item[46] Collins, “Tacit Knowledge, Trust, and the Q of Sapphire.”
\end{footnotes}
\end{footnotesize}
done . . . Step number one, if you want quality, you have to write down what you did in a way that’s unambiguous. Step number two, you have to collect and aggregate all the data on how that performs from a variety of instruments and formats into one place. . . That never happens.”

According to Gardner, many scientists do not have the time to evaluate laboratory contingencies systematically when they arise. Instead of identifying all of the variables that contribute to experimental success, evaluating their impact on the experiment systematically, and modifying protocols as appropriate, many scientists assume the most obvious variable—such as the quality of reagents or the humidity of the laboratory—is responsible for an experiment’s failure. Because scientists do not have the time to evaluate contingencies systematically, Gardner contends, they are not able to learn what is causing the contingency, and key variables that contribute to experimental success and that could be codified remain unrecorded.

During my interview, Gardner, who has been involved in the synthetic biology community by serving as a member of the European Union Scientific Committee’s Working Group on Synthetic Biology, mentioned that gene synthesis firms are very attentive to the problems of reproducibility and are better at ensuring that quality is reproduced (that is, better at making sure their employees achieve consistent results when

---

47 Gardner, interview. Gardner uses the term “quality” in place of “irreproducibility” because irreproducibility is a contentious term. Quality, in Gardner’s opinion, is obtained when a scientists expects a tool or protocol to perform a particular function and the scientist’s expectation is met the majority of the time.

48 J. Webb, “Bringing an End.”

49 Gardner, interview.

50 Ibid.
they follow SOPs) than many other fields. The fact that synthesis firms still encounter difficulties in spite of this rigorous optimization process is telling because it highlights the inherent limitations of working with DNA and microorganisms, which are unpredictable and sensitive to their environments.

While the reproducibility initiatives might achieve some level of success in codifying additional knowledge, STS research shows that not all knowledge that is critical to an experiment’s success can be codified. For this reason, any attempt to improve reproducibility will face challenges because some tacit skills and knowledge cannot be easily explained in writing. These types of knowledge include motor skills that are essential in biology, or knowledge that even the knowledge owner is unaware that he or she possesses. Furthermore, just because an experiment yields inconsistent results does not mean that it cannot be reproduced at all, or that it is not based on credible science.

Despite the fact that scientists often achieve inconsistent results when creating HeLa cell-free extracts or efficiently transplanting the *M. mycoides* genome, the poliovirus and *M. mycoides* experiments can be reproduced; however, researchers might not achieve the same results every time they carry out the protocol. The reproducibility initiatives described above have only attempted to reproduce each experiment once or twice, and have already noted that the original protocols for most experiments were insufficient. Additionally, the STS and organizational theory literature show that having physical access to more experienced scientists is essential for transferring knowledge. For example, biotechnology firms who have access to the inventor of a technology are more

---

51 Gardner, interview.
successful in translating those inventions into commercial products.\textsuperscript{52} If the replicating scientists had physical access to studies’ original authors, they may have been able to replicate their experiments.

These problems with irreproducibility clearly demonstrate that protocols and materials are insufficient for replicating past work, and that tacit knowledge plays a crucial role in experimental success. However, they also suggest that some of this hidden knowledge could be made explicit if scientists had the time and incentive to test out variables of success systematically, and then codify their findings.

\textbf{Finding 3: Because synthesizing genes and using protocols is difficult, researchers rely on tacit knowledge to overcome experimental contingencies}

As mentioned earlier in this chapter, my interviews reinforce previous findings that tacit knowledge is essential to successfully executing life sciences research. Indeed, virtually every aspect of synthetic genomics relies on human knowledge and skills, some of which is tacit. Gene synthesis firms, for example, use the tacit knowledge of their more experienced researchers to resolve difficulties with problematic sequences and to modify or program automated instrumentation and design software. They also use knowledge to perform tasks that are not automated or are only semi-automated, to create and modify SOPs, and even to determine which tasks are best performed by robots. JCVI researchers rely more heavily on tacit knowledge because they are conducting novel research, which often forces them to develop new methods. Indeed, JCVI researchers’ tacit knowledge was a crucial factor in the success of the \textit{M. mycoides} experiment. For example, Dan

\textsuperscript{52} Spinardi, “Defense Technology Enterprises” (see chap. 3, n. 86); Collins, “Tacit Knowledge, Trust, and the Q of Sapphire”; and Zucker, Darby, and Brewer, “Intellectual Human Capital” (see chap. 3, n. 86).
Gibson’s knowledge of genome assembly, which he acquired during the *M. genitalium* experiment, helped him assemble the *M. mycoides* genome. Similarly, Carole Lartigue’s knowledge of the *Mycoplasma* species helped her develop the protocol for genome transplantation. In addition, some of the knowledge that JCVI researchers possess is specialized. For example, Gibson is more adept at genome assembly while Lartigue is more adept at transplantation. Neither individual has all of the knowledge or skills required to perform the entire *M. mycoides* experiment. Moreover, some of the tacit knowledge that JCVI researchers possess has a physical embodiment, requiring them to determine through practice what physical movements are “gentle.” This is consistent with previous findings that having “lab hands” and other tacit motor skills are important to experimental success.⁵³

Even highly routinized procedures, such as PCR, require tacit skills and the involvement of a community of scientists, as seen in the numerous requests on online forums asking for help with PCR. This is also true for procedures that have only recently been routinized, like the Gibson Assembly™ method. The findings noted above on irreproducibility reinforce the importance of tacit knowledge in the sciences. That so many peer-reviewed experiments cannot be reproduced demonstrates that the methods sections of scientific articles do not and often cannot make explicit all of their authors’ tacit knowledge and skills. This is consistent with Collins’ finding that researchers are often not aware of what they know or cannot explain what they know, and therefore

⁵³ Doing, “‘Lab Hands’ and the ‘Scarlet O’” (see chap. 4, n. 74).
cannot easily articulate everything that knowledge recipients need to know in protocols.\textsuperscript{54} This observation also applies to verbal training, in which the knowledge donor who is attempting to transfer knowledge to a recipient (trainee) cannot do so because he or she does not know all of the information the recipient needs to carry out the experiment, or cannot express with words what needs to be done.

As noted earlier, Riffyn’s Gardner believes that tacit knowledge only exists because scientists do not have the time or inclination to make all of their knowledge explicit.\textsuperscript{55} According to Gardner, scientists who have tacit knowledge have detected that they have knowledge and have not yet translated that knowledge into an explicit form, but this will happen over time if scientists are given the tools or incentives to do it.\textsuperscript{56} However, Gardner’s belief is based on the assumption that there is only one form of tacit knowledge; namely, “concealed knowledge,” which Collins defined as knowledge that could be verbalized but remains tacit.\textsuperscript{57} However, Collins shows that tacit knowledge is multidimensional, taking on many forms.\textsuperscript{58} Although some laboratory knowledge could indeed be made explicit, many of these other types of tacit knowledge and skills remain difficult to articulate and make science more of an art. As shown by Collins, transferring and acquiring of this kind of tacit knowledge requires direct interaction between scientists: as they work in the laboratory together, one passes on his skills to the other.

\textsuperscript{54} Collins, “Tacit Knowledge, Trust, and the Q of Sapphire.”
\textsuperscript{55} Gardner, interview.
\textsuperscript{56} Ibid.
\textsuperscript{57} Collins, “Tacit Knowledge, Trust, and the Q of Sapphire.”
\textsuperscript{58} Ibid.
This unarticulated art in science is also what explains why, after 30 years of widespread use, PCR remains a problem for many researchers.

Much of the information about contingencies in the sciences is available in the open literature from the STS field, media interviews with scientists, and online forums that offer troubleshooting services for scientists. However, the majority of threat assessments on biotechnologies fail to consider this rich information, instead preferring to reduce technologies to their material and explicit component while ignoring the social context in which they are used. By omitting this crucial evidence, the threat assessments provide an incomplete characterization of terrorists’ capability to carry out their intent to develop biological agents.

**Finding 4: Most assessments have overstated the threat of synthetic genomics**

Given the above findings, it is clear that most of the mainstream threat assessments overstate the threat posed by synthetic genomics in the near term. These assessments voice concerns that terrorists or other malefactors could use synthetic genomics and synthetic biology to create pathogens. According to the assessments, terrorists could accomplish this in one of two ways: (1) by synthetizing the entire pathogen genome using commercially available, automated synthesizers or (2) by ordering individual pathogen genes from gene synthesis firms and then assembling the genes into a complete genome.

As demonstrated in chapter 4, it would be difficult for malefactors to synthesize pathogen genes or genomes themselves because they would encounter difficulties

---

59 Relman, “Working Group on Synthetic Genomics” (see chap. 2, n. 115).
synthesizing, assembling, and cloning problematic or longer sequences. They may also have trouble verifying that the sequences they synthesized were correct. Chapter 5 demonstrates that malefactors would have additional problems if, instead of synthesizing genes themselves, they decided to order them from synthesis firms. First, they would have problems combining the genes into larger pieces because large pieces of DNA are more likely to behave in unpredictable and undesirable ways. Although commercial reagent kits for creating larger gene constructs exist, such as those that are based on the Gibson Assembly™ method, they are incomplete representations of the expertise required and therefore can still be difficult to use. Second, even if terrorists succeeded in synthesizing an entire pathogen genome, they would have problems transplanting that genome into a host cell using synthetic biology methods.

While synthetic genomics greatly simplifies most scientific experiments, some believe that terrorists would encounter fewer difficulties and possibly achieve greater success if they created biological agents using natural agents. For example, one representative from a gene synthesis firm asserted that “There are much faster and quicker ways to make things when you can just isolate them from nature and do some standard genetic engineering...” because “generating new organisms is a huge undertaking.” However, Ben Ouagrham-Gormley’s research of past biological weapons programs shows that even working with these natural agents is unpredictable and time-consuming. For example, scientists in a research facility in Obolensk, Russia, tried and

\footnote{Ch1b0, “Help!”}

\footnote{Company A, interview, May 20, 2011.}
failed to produce an antibiotic-resistance form of the bacterium that causes plague. Thus, working with natural agents would not make it easier for a terrorist to create a pathogen.

If terrorists or other malefactors wish to create a pathogen using synthetic genomics and synthetic biology, they would be more likely to succeed if they had physical access to individuals with expertise in gene synthesis, genome assembly, and genome transplantation. These experienced individuals would require access to a stable working environment, which would allow them to learn to work together and adapt their tacit knowledge to the new setting without distraction. Nation states are more likely to have access to researchers with this expertise than terrorist or criminal organizations are, and are more likely to provide a stable work environment for the experts. However, historical evidence shows that, despite their vast personnel and material resources, many former biological weapons programs encountered tremendous difficulties and were unsuccessful.

### Additional Considerations

The four findings above are reinforced by historical evidence. To date, despite their clear intent to do so, no terrorist group or nation has exploited advanced biotechnologies to produce biological agents or weapons. Studies of several nations’ former biological weapons programs show that most of these programs failed to develop effective weapons, in spite of using natural agents and having access to protocols,

---

63 Ibid.
equipment, and the tacit knowledge of more experienced researchers. Terrorists who had access to tacit knowledge, materials, and protocols also failed to produce working agents. For example, the supposed perpetrator of the 2001 anthrax letter attacks was only able to produce a low-grade powder, and Japan’s Aum Shinrikyo cult failed during all stages of their attempts to produce weapons that were based on anthrax and botulinum toxin. This demonstrates that even work with natural agents is not as simple as some gene synthesis scientists today assume. Even if terrorists were able to succeed in producing pathogens, they would face the additional challenge of converting the resultant pathogen into a form that can be produced in larger quantities and delivered and disseminated without loss of its pathogenic properties—problems that frustrated U.S. and Soviet efforts in developing effective biological weapons.

Thus, it is clear that threat assessments overstate the dual-use threat of synthetic genomics because they place great emphasis on terrorists’ intent to develop biological agents while assuming that their capability to do so merely involves acquiring materials and protocols. In adopting these assumptions, security researcher Ben Ouagrham-Gormley believes the assessments inadvertently divert attention away from the variable that has greater impact on their capability to develop biological agents: knowledge.

The following chapter describes my recommendations for improving threat assessments and scientific protocols and mitigating the risk that synthetic genomics and synthetic biology will be misused by terrorists or other malefactors.

---

64 Ibid.
65 Ibid., 3–5.
66 Ibid.
67 Ibid., 12.
CHAPTER 7
CONCLUSIONS AND FURTHER AREAS OF STUDY

My research indicates that it would be difficult for terrorists to synthesize and assemble pathogen genomes because gene synthesis firms and researchers at prominent institutions like the JCVI encounter difficulties performing these tasks and because the difficulties these experts encounter are overcome using tacit knowledge and skills that are acquired through years of hands-on practice. This is because explicit forms of knowledge, like protocols and procedures, do not contain all of the knowledge that their authors possess and therefore require significant interpretation and modification by end-users. STS research indicates that this tacit knowledge is difficult to acquire and transfer outside of certain contexts, further demonstrating that terrorists are unlikely to successfully use synthetic genomics and synthetic biology to create pathogens.

These findings run counter to many current threat assessments of biotechnologies, which assume that a malefactor’s capability to develop biological agents using synthetic genomics merely requires access to technology and protocols; skill level, if considered at all, is only discussed in vague terms and generally not factored into the assessment. Current threat assessment methodologies, however, are inadequate because they fail to examine the technical skills required to use biotechnologies. Thus, these threat assessment methodologies should be modified to include technical, sociological, and
organizational details about how scientists develop and use dual-use technologies.

Second, many biosecurity initiatives and nonproliferation policies fail to consider the sociological factors involved in developing and using technologies, focusing instead on materials and technologies. Thus, these initiatives and policies should incorporate a sociotechnical frame. Third, despite encountering many technical hurdles, scientists are nevertheless able to acquire the tacit knowledge they need to carry out an experiment if given the time, resources, and access to relevant expertise. Although it is unlikely—especially if they lack the appropriate expertise and lack access to more experienced researchers—persistent malefactors might eventually acquire the tacit knowledge that would allow them to create a pathogen using synthetic genomics and synthetic biology if they gain access to relevant expertise and have the time and resources to troubleshoot the myriad of issues that are sure to arise. As such, the life sciences community should embrace—and the government should strengthen—measures that are aimed at preventing dual-use technologies such as synthetic genomics and synthetic biology from being misused. Lastly, current initiatives in academia that aim to study the irreproducibility problem will document the causes of irreproducibility in the future and provide a very rich source of data for STS researchers and national security analysts. STS and national security analysts who plan to use this data should engage with the researchers who are spearheading these efforts to ensure they develop mechanisms for capturing sociological and organizational data relating to irreproducibility—not just technical data. These four recommendations are described in greater detail in the following sections.
Recommendation 1: Add credibility to threat assessments by characterizing technologies within a sociotechnical frame

My research confirms what other researchers have already highlighted: current methodologies for evaluating the threat posed by biotechnologies are inadequate because the current methodologies do not adequately examine the technical skills required to use biotechnologies.475 As discussed in chapter 2, many mainstream threat assessments are written to describe the hostile applications of a technology, thus they only consider what is theoretically possible given materials and procedures and pay little attention to the capability required to develop and use the technology. As such, they provide incomplete descriptions of the actual threat stemming from biotechnologies. Thus, threat assessment methodologies should be reevaluated so that they can include nuanced details about how scientists develop and use dual-use technologies.

One major shortcoming of threat assessments on biotechnologies is that they fail to consider the rich STS literature that describes contingencies in life science research. Indeed, STS researchers who evaluated academic and nonacademic threat assessments published since 1999 found that none drew upon or cited STS research, and the phrase “tacit knowledge” did not appear in any of the assessments.476 STS researcher Vogel contends that technologies should be understood as involving knowledge, skills, practices, and tools—not just materials and hardware.477 In other words, national security analysts should try to understand the dual-use technologies within a sociotechnical frame.

---

475 Vogel, Phantom Menace (see chap. 1, n. 11); Ben Ouagrham-Gormley, Barriers to Bioweapons (see chap. 1, n. 11).
476 Marris, Jefferson, and Lentzos, “Negotiating the Dynamics of Uncomfortable Knowledge” (see chap. 2, n. 62).
477 Vogel, Phantom Menace.
that focuses on tacit knowledge, communities of practice, organizational factors, and technical difficulties. Conducting this type of nuanced analysis of a technology’s social context is what Vogel calls adopting a “biosocial frame” for understanding technical work.\textsuperscript{478} According to Vogel, when analysts adopt this frame, they are prompted to understand laboratory practices, identify how knowledge is generated in the laboratory, and identify which organizational factors might contribute to success in the laboratory.\textsuperscript{479}

The easiest way for analysts to obtain the data that Vogel recommends is to interview the researchers who are actually performing the research in question. Analysts should undertake these interviews with the goal of understanding all stages of the experiment, identifying any steps where contingencies arise, and determining what is required to overcome contingencies. This will help them understand what sorts of hurdles the scientists who are developing or using the technologies face, how they overcome those difficulties, which type of knowledge they primarily rely on when conducting the research, how easily their experiments could be applied to different social and physical contexts, and what level of tacit knowledge is involved in using tools and interpreting protocols. Importantly, analysts must ask these questions in the context of understanding what goes into the research; they should not simply ask interview subjects their opinions about how easily malefactors could exploit the technology, since this might provide biased information.

\textsuperscript{478} Vogel, \textit{Phantom Menace}. According to Vogel, the biosocial frame emphasizes how technical knowledge, work, and products are mutually constituted with a given social context. Instead of focusing only on materials and technologies, the biosocial frame involves studying laboratory practices and understanding the laboratory’s organizational structure.

\textsuperscript{479} Vogel, \textit{Phantom Menace}. 

211
Security researcher Ben Ouagrham-Gormley also offers an improved framework for national security analysts for determining whether a terrorist organization’s procurement of materials and expertise will yield biological agents. According to Ben Ouagrham-Gormley, analysts should consider the group’s absorptive capacity (the extent to which it integrates acquired material and expertise into its existing knowledge base), the organizational makeup of the group, and the group’s managerial characteristics.\textsuperscript{480} Evaluating absorptive capacity involves determining whether individuals working with the technologies have expertise that corresponds to the work they are doing. This can be a challenge in the biological sciences, where technical expertise is very specialized. For example, an individual with decades of experience working with bacteria will not necessarily be able to work with viruses because, among other differences, viruses have different growth requirements than bacteria that the individual will need to master. Similarly, an individual whose doctoral research involves identifying proteins that allow the influenza virus to infect human cells might not have any actual knowledge of how to grow influenza in culture, because his or her work may have only dealt with studying the entry mechanisms using less pathogenic, proxy pathogens. Understanding an organization’s makeup will allow analysts to determine what types of knowledge can be created and how it can be used within an organization. According to Ben Ouagrham-Gormley, open, flexible programs that promote communal knowledge are more conducive to knowledge transfer and technical progress.\textsuperscript{481}

\textsuperscript{480} Ben Ouagrham-Gormley, \textit{Barriers to Bioweapons}, 13.
\textsuperscript{481} Ben Ouagrham-Gormley, \textit{Barriers to Bioweapons}.
organization’s management characteristics allows analysts to project outcomes and timelines. Organizations with scientists whose skills are not directly applicable to the experiments they are conducting, a vertical management structure, and no methods for integrating disparate knowledge within the organization will likely make little progress in developing a biological agent. The opposite is true of organizations that are staffed with scientists with relevant knowledge, a horizontal management structure, and ways to integrate disparate knowledge within the organization.482 My interviews with gene synthesis firms and JCVI researchers indicate that the management structure of both organizations was open and collaborative, where employees were encouraged to network and interact with other scientists within and outside of the organization and were given the freedom to “explore condition space” to resolve problems.483

Because terrorists and other malefactors usually operate covertly, analysts who adopt the biosocial frame for assessments likely will not have access to all of the sociotechnical data they need to evaluate, such as data about the management style of a lead scientist in a terrorist group program. In addition, critics might argue that the only true way to quantify or characterize the level and type of expertise or human interaction necessary to produce biological agents or weapons using biotechnology is impossible unless it can be empirically tested (that is, attempted in the laboratory). The Department of Homeland Security’s National Biodefense Analysis and Countermeasures Center (NBACC), created in 2004 to characterize biological threats, may be conducting

482 Ibid., 153–155.
483 Glass, interview (see chap. 5, n. 5).
experiments like these as part of its biodefense work. The NBACC’s Biological Threat Characterization Center (BTCC) conducts studies and laboratory experiments to identify and address gaps in the United States’ understanding of biological threats, including select agent pathogens. While these biodefense experiments could serve as good proxies for studying how malefactors might acquire skills for working with pathogens, conducting the experiments in large quantities would be prohibited under the Biological and Toxin Weapons Convention (BTWC). However, adopting the biosocial frame will provide more complete information than what is provided in many mainstream threat assessments because many of these assessments are essentially technology forecasts, which are notoriously vague and unreliable. By incorporating studies that examine the social context of science of concern into threat assessments, national security analysts can provide crucial insight about technologies that will increase the credibility of their assessments. In cases where terrorist organizations have acquired expertise and material, analysts could potentially improve the specificity of their forecast timelines by determining organizations’ absorptive capacity, organizational makeup, and managerial

486 A 2007 Congressional Research Service report indicated that arms control experts had raised concern that some of the work conducted at the BTCC could be interpreted as violating the Biological and Toxin Weapons Convention [See Shea, “National Biodefense Analysis and Countermeasures Center.”]. To allay these fears, the DHS has implemented measures to ensure the NBACC complies with arms control agreements [See Stacy Okutani, “The U.S. Threat Assessment Process and Oversight for Biological Research at NBACC,” Center for International and Security Studies at Maryland, December 2007, http://drum.lib.umd.edu/bitstream/handle/1903/15982/u.s_threat_assessment_process_and_oversight_of_biological_research_at_nbacc.pdf.]
characteristics because these factors provide insight about how quickly an organization could develop biological agents. In addition, adopting the biosocial frame would provide data in less time than it would take to conduct a proxy experiment. Finally, the biosocial frame will also withstand international scrutiny because, unlike proxy experiments, it would not give other nations the impression that the United States is violating arms control treaties.

**Recommendation 2: Incorporate a sociotechnical frame into national biosecurity initiatives and international nonproliferation policies**

In addition to analysts who generate threat assessments on dual-use technologies, international nonproliferation policies and federal advisory groups on dual-use technologies—both of which typically focus on preventing access to or regulating materials—should also incorporate the sociotechnical frame into their deliberations. The BTWC, for example, should encourage states parties to consider the role of tacit knowledge on developments in science and technology. During the Seventh Review Conference 2012, states parties agreed that the Eighth Review Conference should take into account new scientific and technological developments that relate to the BTWC.\(^{488}\) At a meeting of experts held in 2015, members recommended that future discussions between states parties consider the role of tacit knowledge in science and technological development because tacit knowledge can modulate the risks associated with biological weapons production and proliferation. That is, the lack of tacit knowledge can decrease

the risk that a proliferator would be able to produce biological weapons using only materials and protocols, whereas the converse is true for proliferators who have access to tacit knowledge. Future BTWC meetings should consider the sociological factors associated with developing and using dual-use technologies, including how tacit knowledge factors in and how its proliferation can be thwarted. This will help states parties determine which technologies warrant concern.

BTWC states parties should also consider tacit knowledge when developing policies for preventing proliferation. By understanding how individuals and organizations acquire and transfer knowledge, BTWC members could develop more effective policies. For example, BTWC members could establish policies that require scientists with knowledge of biological weapons production to be transferred to a different laboratory where they will work in a new setting, with a new team, and with a different organism. This policy would allow the knowledge that scientists possess about biological weapons production to atrophy, reducing the risk of proliferation. This would be an improvement over current approaches, which often allow these scientists to work in the same laboratories where they once conducted biological weapons research, and to work with the same organisms they had once tried to weaponize.

The NSABB should also incorporate the sociotechnical frame into their deliberations, which up to now have focused primarily on the materials necessary to misuse dual-use research and technologies. Incorporating the sociotechnical frame will

help the NSABB identify more appropriate measures for preventing dual-use research from being misused, because it will allow them to factor tacit knowledge and skills into their deliberations. For example, a recent JCVI report that assessed how well the DHHS’ screening guidance for synthetic DNA providers has worked since its release, suggested that the DHHS consider incorporating benchtop gene synthesizers into its screening guidance in the future. Although there is currently only one such synthesizer available, the JCVI anticipated that more synthesizers would enter the market.\textsuperscript{490} To ensure that these technologies are not misused, the JCVI suggested that manufacturers of the synthesizers screen their customer and the gene sequences they would be producing.\textsuperscript{491} While these types of interventions may be warranted, the NSABB should first consider the sociological factors involved in using these benchtop synthesizers, as well as the sociological factors involved in using reagent kits for constructing genes, such as those that are based on the Gibson Assembly\textsuperscript{TM} method. As discussed in previous chapters, although manufacturers claim that their reagent kits, protocols, and equipment are rapid and easy to use, previous STS research and my interviews suggest otherwise. Indeed, online forums for troubleshooting problems in laboratory research are replete with examples of how these technologies must be adapted to suit the needs of the laboratory or experiment in which they are being used. By first considering and evaluating how

\textsuperscript{490} SGI-DNA’s BioXP\textsuperscript{TM} 3200 System is currently the only benchtop gene synthesizer available. Customers who use this system must send the sequences they wish to synthesize to SGI-DNA’s secure website, which then designs and optimizes the sequence. SGI-DNA, an IGSC member, screens these sequences per the IGSC’s Harmonized Screening Protocol before initiating the design and optimization. However, future manufacturers of benchtop DNA synthesizers might not offer a design and optimization service; in these cases, the customer would simply purchase the benchtop synthesizer and begin producing DNA, bypassing sequence screening altogether.\textsuperscript{491} Carter and Friedman, “DNA Synthesis and Biosecurity” (see chap. 2, n. 124).
scientists use these technologies, the NSABB can generate policy recommendations that are more effective at preventing technologies from being misused because those policies will address expertise—a key variable in proliferation—as well as access to material.

**Recommendation 3: Embrace and improve measures to prevent the misuse of synthetic genomics and synthetic biology**

Because scientists are able to acquire the tacit knowledge they need to carry out an experiment despite the technical hurdles they face, there is the possibility that persistent, patient malefactors could eventually acquire the tacit knowledge that would enable them to create a pathogen using synthetic genomics and synthetic biology. As such, members of the life sciences community should embrace measures that seek to prevent this misuse. Specifically, the government and gene synthesis firms should reinforce industry norms for sequence screening and also improve compliance with current screening guidelines by providing free tools to help firms screen their orders and customers. In addition, the U.S. government should continue to insist that any research organization it funds conduct regular institutional reviews of all of their research projects—including those that are not funded by the government—to identify dual-use research of concern (DURC) and develop methods for mitigating risks associated with it.

Current guidance for sequence screening and institutional review of DURC are good starting points for preventing the misuse of synthetic genomics, synthetic biology, and other DURC; however, there is room for improvement. For example, some synthetic gene providers do not screen orders to identify sequences or customers of concern. This is because complying with existing guidance can be burdensome, requiring a human with
specialized knowledge to identify sequences or concern and, in some cases, verify the legitimacy of customers’ research. A recent report released by the JCVI estimated that approximately five percent of orders that International Gene Synthesis Consortium (IGSC) members fill require investigation, and that these investigations can take up to 90 minutes to resolve. The labor costs associated with these investigations can place a large financial burden on synthesis firms. Declining costs for gene synthesis and competition from international providers of synthetic genes might increase this burden in the future. Therefore, the U.S. government should undertake efforts to ensure greater compliance with screening guidance, as well as decrease the burdens of screening. For example, the government could adopt the JCVI’s recommendation that it require its grantees and contractors to purchase genes only from companies that comply with the DHHS screening framework. Additionally, the government could adopt the JCVI’s recommendation to provide a free database containing sequences of concern that gene synthesis firms can use during order screening. The JCVI recommended that the database be updated frequently to include new threats and that the database prioritize sequences that a scientist with a moderate skill set could render pathogenic. These two measures would help to reinforce biosecurity norms in the industry while also making it easier for companies to comply with existing guidance.

As described in chapter 2, the government requires that its employees, contractors, and grantees who work with synthetic genes or conduct dual-use research of concern (DURC) conduct regular institutional reviews to identify their DURC and

492 Ibid.
implement measures to reduce the risk that the materials or technologies will be misused. However, IRBs might lack the knowledge to identify DURC and they might not implement risk mitigation measures in such a way as to prevent its misuse. These factors point to the potential need for increased government involvement in helping laboratories both identify DURC and develop appropriate risk mitigation measures. In addition, current policies only require institutional reviews for government agencies or organizations that receive government funds. The government should encourage commercial firms who conduct DURC but do not receive government funding to implement measures to prevent materials and knowledge from being misused.

Lastly, IRBs primarily address ways to mitigate the misuse of dual-use materials, not knowledge. This indicates that government training on biosafety should incorporate the sociotechnical frame to ensure that IRBs control the knowledge they disseminate that is related to DURC. For example, IRBs might prohibit a researcher from publishing in the Journal of Visualized Experiments if they know the researcher’s experiment does not require knowledge or skills that are especially difficult to convey even through video demonstration, such as motor skills. By determining the sociological factors involved in conducting experiments, IRBs can identify ways to prevent the knowledge associated with DURC from being misused.

**Recommendation 4: Engage with current reproducibility initiatives to ensure data that is relevant to STS and national security researchers is collected**

My research adds to the body of literature showing that published scientific protocols can be difficult to reproduce. Although my research and previous STS studies
show that difficulties are due to the lack of tacit knowledge embodied in protocols and other explicit forms of knowledge, current initiatives in academia that study the irreproducibility problem, such as the Reproducibility Project: Cancer Biology, will document more causes of irreproducibility in the future. These initiatives will provide a wealth of information that STS researchers and national security analysts could use to understand the limitations of explicit knowledge and determine with more accuracy how easy or difficult it would be for malefactors to replicate life sciences experiments. Therefore, STS and national security analysts who plan to use this data should engage with the researchers who are spearheading these efforts to ensure the efforts capture metrics that are important to the STS and national security communities, such as how frequently crucial information was omitted from scientific articles and why it was omitted, or to what extent laboratory reagents or tools affected reproducibility. In addition to ensuring that the initiatives provide meaningful data, the STS community in particular has much to offer the reproducibility community, in terms of helping its members understand what factors might influence a study’s reproducibility and what data fields might be important to include in software tools.

**Conclusions**

This dissertation has shown the technical difficulties that terrorists and other malefactors would encounter if they tried to synthesize pathogen genomes themselves. Because there is evidence that synthesizing large genes is becoming less difficult, future studies should characterize the ease with which scientists use commercial large-gene and genome assembly techniques such as the Gibson Assembly™ method. These studies
could elucidate whether the commercially available techniques have appreciably decreased difficulties associated with assembling large genes and entire genomes. One way to study whether assembling large genes and genomes has become appreciably less difficult is to conduct a case study of the Sc2.0 project, which aims to synthesize the entire 14 Mb genome of *Saccharomyces cerevisiae* (that is, Sc), a yeast with a genome that is much larger than that of *M. mycoides*.\footnote{“Goals,” *Synthetic Yeast 2.0* (blog), Science Across Virtual Institutes, Accessed November 6, 2015, http://syntheticyeast.org/sc2-0/goals/} The project involves hundreds of researchers from around the world with varying skill sets—from high school students to scientists with decades of experience. This case study would provide valuable data on how scientists with different backgrounds and levels of experience fare in synthesizing genes, assembling large genomes, and whether geographic location or other factors affect success.

In addition to case studies on large-gene synthesis, future STS work should focus on reproducibility initiatives, such as the Open Science Framework’s Reproducibility Project: Cancer Biology, because conducting case studies of some irreproducible experiments will help STS researchers further characterize the difficulties involved in conducting life sciences research and conveying tacit knowledge.

While my research examined the difficulties that terrorists might encounter in using synthetic genomics and synthetic biology to create a pathogen, future studies should examine the difficulties a nation state might encounter when conducting similar experiments, since states are more likely to have access to experts in synthetic genomics.
and to provide a stable work environment for these experts. Previous research of state biological weapons programs shows that, despite their vast resources, scientists in state programs encountered tremendous difficulties when working with natural pathogens and using traditional genetic engineering techniques.\textsuperscript{494} Future studies should determine whether these findings hold true for synthetic genomics. This will provide valuable insight about the adequacy of current nonproliferation policies to address threats from synthetic genomics and synthetic biology.

Lastly, while my research examined the difficulties associated with whole-genome synthesis, future studies should examine the difficulties terrorists or other malefactors might encounter when using other methods to modify or create pathogens, which might be less challenging than synthesizing a pathogen’s entire genome.

\textsuperscript{494} Ibid.
APPENDIX A: SELECT AGENTS AND TOXINS LIST

§ 73.3 HHS select agents and toxins

(a) Except for exclusions under paragraphs (d) and (e) of this section, the HHS Secretary has determined that the biological agents and toxins listed in this section have the potential to pose a severe threat to public health and safety. The select agents and toxins marked with an asterisk (*) are designated as Tier 1 select agents and toxins and are subject to additional requirements as listed in this part.

(b) HHS select agents and toxins:

- Abrin
- Botulinum neurotoxins*
- Botulinum neurotoxin producing species of *Clostridium*

Conotoxins (short, paralytic alpha conotoxins containing the following amino acid sequence $X_1CCX_2PACGX_3X_4X_5X_6CX_7$)

- *Coxiella burnetii*
- Crimean-Congo haemorrhagic fever virus
- Diacetoxycircpenol
- Eastern Equine Encephalitis virus
- Ebola virus*
• *Francisella tularensis*
• Lassa fever virus
• Lujo virus
• Marburg virus*
• Monkeypox virus
• Reconstructed replication competent forms of the 1918 pandemic influenza virus containing any portion of the coding regions of all eight gene segments (Reconstructed 1918 Influenza virus)
• Ricin
• *Rickettsia prowazekii*
• *SARS-associated coronavirus (SARS-CoV)*
• Saxitoxin
• South American Haemorrhagic Fever viruses:
  o Chapare
  o Guanarito
  o Junin
  o Machupo
  o Sabia
• Staphylococcal enterotoxins (subtypes A-E)
• T-2 toxin
• Tetrodotoxin
• Tick-borne encephalitis virus
• Far Eastern subtype
• Siberian subtype
• Kyasanur Forest disease virus
• Omsk haemorrhagic fever virus
• Variola major virus (Smallpox virus)*
• Variola minor virus (Alastrim)*
• *Yersinia pestis*

(c) Genetic Elements, Recombinant and/or Synthetic Nucleic Acids, and Recombinant and/or Synthetic Organisms:

(1) Nucleic acids that can produce infectious forms of any of the select agent viruses listed in paragraph (b) of this section.

(2) Recombinant and/or Synthetic nucleic acids that encode for the functional form(s) of any of the toxins listed in paragraph (b) of this section if the nucleic acids:

(i) Can be expressed in vivo or in vitro, or

(ii) Are in a vector or recombinant host genome and can be expressed in vivo or in vitro.

(3) HHS select agents and toxins listed in paragraph (b) of this section that have been genetically modified.

(d) HHS select agents or toxins that meet any of the following criteria are excluded from the requirements of this part:
(1) Any HHS select agent or toxin that is in its naturally occurring environment provided the select agent or toxin has not been intentionally introduced, cultivated, collected, or otherwise extracted from its natural source.

(2) Non-viable HHS select agents or nonfunctional HHS toxins.

(3) Except as required in §73.16(l), the aggregate amount of the toxin under the control of a principal investigator, treating physician or veterinarian, or commercial manufacturer or distributor does not, at any time, exceed the following amounts: 100 mg of Abrin; 0.5 mg of Botulinum neurotoxins; 100 mg of Conotoxins (Short, paralytic alpha conotoxins containing the following amino acid sequence X_1CCX_2PACGX_3X_4X_5X_6CX_7); 1,000 mg of Diacetoxyispirenol; 100 mg of Ricin; 100 mg of Saxitoxin; 5 mg of Staphylococcal enterotoxins (subtypes A-E); 1,000 mg of T-2 toxin; or 100 mg of Tetrodotoxin.

   (i) The amounts are transferred only after the transferor uses due diligence and documents that the recipient has a legitimate need (i.e., reasonably justified by a prophylactic, protective, bona fide research, or other peaceful purpose) to handle or use such toxins. Notwithstanding the provisions of paragraph (d) of this section, the HHS Secretary retains the authority to, without prior notification, inspect and copy or request the submission of the due diligence documentation to the CDC.

   (ii) Reports to CDC if they detect a known or suspected violation of Federal law or become aware of suspicious activity related to a toxin listed in this part.
(4) An animal inoculated with or exposed to an HHS select toxin.

(5) Any South American genotypes of Eastern Equine Encephalitis Virus and any West African Clade of Monkeypox virus provided that the individual or entity can identify that the agent is within the exclusion category.

(e) An attenuated strain of a select agent or a select toxin modified to be less potent or toxic may be excluded from the requirements of this part based upon a determination by the HHS Secretary that the attenuated strain or modified toxin does not pose a severe threat to public health and safety.

(1) To apply for exclusion, an individual or entity must submit a written request and supporting scientific information. A written decision granting or denying the request will be issued. An exclusion will be effective upon notification to the applicant. Exclusions will be listed on the National Select Agent Registry Web site at http://www.selectagents.gov/.

(2) If an excluded attenuated strain or modified toxin is subjected to any manipulation that restores or enhances its virulence or toxic activity, the resulting select agent or toxin will be subject to the requirements of this part.

(f) Any HHS select agent or toxin seized by a Federal law enforcement agency will be excluded from the requirements of this part during the period between seizure of the select agent or toxin and the transfer or destruction of such agent or toxin provided that:

(1) As soon as practicable, the Federal law enforcement agency transfers the seized select agent or toxin to an entity eligible to receive such agent or toxin or destroys the agent or toxin by a recognized sterilization or inactivation process,
(2) The Federal law enforcement agency safeguards and secures the seized select agent or toxin against theft, loss, or release, and reports any theft, loss, or release of such agent or toxin, and

(3) The Federal law enforcement agency reports the seizure of the select agent or toxin to CDC or APHIS.

(i) The seizure of Botulinum neurotoxins, Botulinum neurotoxin producing species of *Clostridium*, Ebola viruses, *Francisella tularensis*, Marburg virus, Variola major virus (Smallpox virus), Variola minor (Alastrim), or *Yersinia pestis* must be reported within 24 hours by telephone, facsimile, or e-mail. This report must be followed by submission of APHIS/CDC Form 4 within seven calendar days after seizure of the select agent or toxin.

(ii) For all other HHS select agents or toxins, APHIS/CDC Form 4 must be submitted within seven calendar days after seizure of the agent or toxin.

(iii) A copy of APHIS/CDC Form 4 must be maintained for three years.

(4) The Federal law enforcement agency reports the final disposition of the select agent or toxin by submission of APHIS/CDC Form 4. A copy of the completed form must be maintained for three years.

APPENDIX B: AUSTRALIA GROUP LIST OF HUMAN AND ANIMAL PATHOGENS AND TOXINS FOR EXPORT CONTROL [1]

Viruses

African horse sickness virus
African swine fever virus
Andes virus
Avian influenza virus [2]
Bluetongue virus
Chapare virus
Chikungunya virus
Choclo virus
Classical swine fever virus (Hog cholera virus)
Crimean-Congo hemorrhagic fever virus
Dengue virus
Dobrava-Belgrade virus
Eastern equine encephalitis virus
Ebolavirus: all members of the Ebolavirus genus
Foot-and-mouth disease virus
Goatpox virus
Guanarito virus
Hantaan virus
Hendra virus (Equine morbillivirus)
Japanese encephalitis virus
Junin virus
Kyasanur Forest disease virus
Laguna Negra virus
Lassa virus
Louping ill virus
Lujo virus
Lumpy skin disease virus
Lymphocytic choriomeningitis virus
Machupo virus
Marburgvirus: all members of the Marburgvirus genus
Monkeypox virus
Murray Valley encephalitis virus
Newcastle disease virus
Nipah virus
Omsk hemorrhagic fever virus
Oropouche virus
Peste-des-petits-ruminants virus
Porcine Teschovirus
Powassan virus
Rabies virus and other members of the Lyssavirus genus
Reconstructed 1918 influenza virus
Rift Valley fever virus
Rinderpest virus
Rocio virus
Sabia virus
Seoul virus
Severe acute respiratory syndrome-related coronavirus (SARS-related coronavirus)
Sheeppox virus
Sin Nombre virus
St. Louis encephalitis virus
Suid herpesvirus 1 (Pseudorabies virus; Aujeszky's disease)
Swine vesicular disease virus
Tick-borne encephalitis virus (Far Eastern subtype)
Variola virus
Venezuelan equine encephalitis virus
Vesicular stomatitis virus
Western equine encephalitis virus
Yellow fever virus

**Bacteria:**

*Bacillus anthracis*
Brucella abortus
Brucella melitensis
Brucella suis
Burkholderia mallei (Pseudomonas mallei)
Burkholderia pseudomallei (Pseudomonas pseudomallei)
Chlamyphila psittaci (formerly known as Chlamydia psittaci)
Clostridium argentinense (formerly known as Clostridium botulinum Type G), botulinum neurotoxin producing strains
Clostridium baratii, botulinum neurotoxin producing strains
Clostridium botulinum
Clostridium butyricum, botulinum neurotoxin producing strains
Clostridium perfringens, epsilon toxin producing types [3]
Coxiella burnetii
Francisella tularensis
Mycoplasma capricolum subspecies i (“strain F38”)
Mycoplasma mycoides subspecies i SC (small colony)
Rickettsia prowazekii
Salmonella typhi
Shiga toxin producing Escherichia coli (STEC) of serogroups O26, O45, O103, O104, O111, O121, O145, O157, and other shiga toxin producing serogroups [4]
Shigella dysenteriae
Vibrio cholerae
Yersinia pestis

Toxins as follows and subunits thereof: [5]

Abrin
Aflatoxins
Botulinum toxins[6]
Cholera toxin

Clostridium perfringens alpha, beta 1, beta 2, epsilon and iota toxins
Conotoxin [6]
Diacetoxyscirpenol toxin
HT-2 toxin
Microcystin (Cyanoginosin)
Modeccin toxin
Ricin
Saxitoxin
Shiga toxin

Staphylococcus aureus enterotoxins, hemolysin alpha toxin, and toxic shock syndrome toxin (formerly known as Staphylococcus enterotoxin F)

T-2 toxin
Tetrodotoxin
Verotoxin and shiga-like ribosome inactivating proteins

Viscum Album Lectin 1 (Viscumin)
Volkensin toxin

234
**Fungi**

*Coccidioides immitis*

*Coccidioides posadasii*

[1] An agent/pathogen is covered by this list except when it is in the form of a vaccine. A vaccine is a medicinal product in a pharmaceutical formulation licensed by, or having marketing or clinical trial authorization from, the regulatory authorities of either the country of manufacture or of use, which is intended to stimulate a protective immunological response in humans or animals in order to prevent disease in those to whom or to which it is administered.

Biological agents and pathogens are controlled when they are an isolated live culture of a pathogen agent, or a preparation of a toxin agent which has been isolated or extracted from any source, or material including living material which has been deliberately inoculated or contaminated with the agent. Isolated live cultures of a pathogen agent include live cultures in dormant form or in dried preparations, whether the agent is natural, enhanced or modified.

[2] This includes only those Avian influenza viruses of high pathogenicity as defined by the World Organization for Animal Health (OIE), the European Union (EU), or competent national regulatory bodies.

[3] It is understood that limiting this control to epsilon toxin-producing strains of *Clostridium perfringens* therefore exempts from control the transfer of other *Clostridium perfringens* strains to be used as positive control cultures for food testing and quality control.
[4] Shiga toxin producing *Escherichia coli* (STEC) is also known as enterohemorrhagic *E. coli* (EHEC) or verocytotoxin producing *E. coli* (VTEC).

[5] Excluding immunotoxins

[6] Excluding botulinum toxins and conotoxins in product form meeting all of the following criteria:

- are pharmaceutical formulations designed for testing and human administration in the treatment of medical conditions;
- are pre-packaged for distribution as clinical or medical products; and
- are authorized by a state authority to be marketed as clinical or medical products

**Warning List [7]**

*Clostridium tetani* [8]

*Legionella pneumophila*

*Yersinia pseudotuberculosis*

Other strains of *Clostridium* species that produce botulinum neurotoxin [9]

**Fungi**

*Fusarium langsethiae*

*Fusarium sporotrichioides*

[7] Biological agents are controlled when they are an isolated live culture of a pathogen agent, or a preparation of a toxin agent which has been isolated or extracted from any source, or material including living material which has been deliberately inoculated or contaminated with the agent. Isolated live cultures of a pathogen agent
include live cultures in dormant form or in dried preparations, whether the agent is natural, enhanced or modified.

An agent is covered by this list except when it is in the form of a vaccine. A vaccine is a medicinal product in a pharmaceutical formulation licensed by, or having marketing or clinical trial authorization from, the regulatory authorities of either the country of manufacture or of use, which is intended to stimulate a protective immunological response in humans or animals in order to prevent disease in those to whom or to which it is administered.

[8] The Australia Group recognizes that this organism is ubiquitous, but, as it has been acquired in the past as part of biological warfare programs, it is worthy of special caution.

[9] It is the intent of Australia Group members to add to the control list strains of species of *Clostridium* identified as producing botulinum neurotoxin.

**Genetic elements and genetically modified organisms:**

- Genetic elements that contain nucleic acid sequences associated with the pathogenicity of any of the microorganisms in the list.
- Genetic elements that contain nucleic acid sequences coding for any of the toxins in the list, or for their sub-units.
- Genetically modified organisms that contain nucleic acid sequences associated with the pathogenicity of any of the microorganisms in the list.
- Genetically modified organisms that contain nucleic acid sequences coding for any of the toxins in the list or for their sub-units.
**Technical note:**

Genetically modified organisms includes organisms in which the genetic material (nucleic acid sequences) has been altered in a way that does not occur naturally by mating and/or natural recombination, and encompasses those produced artificially in whole or in part.

Genetic elements include inter alia chromosomes, genomes, plasmids, transposons, and vectors whether genetically modified or unmodified, or chemically synthesized in whole or in part.

Nucleic acid sequences associated with the pathogenicity of any of the microorganisms in the list means any sequence specific to the relevant listed micro-organism:

- that in itself or through its transcribed or translated products represents a significant hazard to human, animal or plant health; or

- that is known to enhance the ability of a listed micro-organism, or any other organism into which it may be inserted or otherwise integrated, to cause serious harm to human, animal or plant health. These controls do not apply to nucleic acid sequences associated with the pathogenicity of enterohemorrhagic *Escherichia coli*, serotype O157 and other verotoxin producing strains, other than those coding for the verotoxin, or for its sub-units.
APPENDIX C: IASB CODE OF CONDUCT

1. Preamble


Preamble

The field of Synthetic Biology is gaining momentum in the academic and commercial world and evolving rapidly. In parallel, a market for Synthetic Biology products and services has developed and grown rapidly over the past ten years. The International Association Synthetic Biology represents a number of companies and organizations with a stake in Synthetic Biology, for instance as providers of double-stranded recombinant DNA synthesis (hereinafter “gene synthesis”) or bioinformatics products. IASB has created this Code of Conduct in order to secure the foundations of this fledgling field against abuse and to bring Synthetic Biology to its full potential. It is aimed at all providers of gene synthesis services. The most fundamental tools for the design of Synthetic Biology applications are synthetic genes and their intrinsic features of freedom of design and artificial biological function. This Code of Conduct helps companies that provide DNA synthesis services and products and academic and public institutions that practice DNA synthesis to conduct their business in a sensible and responsible way. Declaration: The Undersigned herewith declare that they are in full agreement with the need for a safe and responsible use of synthetic DNA. They strictly
follow all regulations and international standards designed to safeguard against intentional or unintentional abuse of synthetic DNA.

2. **General Considerations**

   Synthetic Biology provides the means to accelerate the assembly of complex biological networks and to rapidly create biological entities with new properties. These powers will undoubtedly lead to a number of beneficial developments such as sustainable biofuels, new therapeutics, and biodegradable plastics. However, the efficiency and potential power of Synthetic Biology can also create the risk of abuse. Through rapid DNA synthesis, biorisk-associated genes such as toxin genes or virulence factors become accessible to a large number of users. In order to contain the risks of Synthetic Biology and to protect the field against misuse, the Undersigned have adopted this Code of Conduct which provides guidelines for safe, secure, and responsible commercial or non-commercial DNA synthesis. One important consideration of any regulation for biosafety and biosecurity is the freedom of research: A lot of beneficial developments would be impossible without the freedom to explore organisms and genes that bear a certain environmental or health risk. It is our conviction that such a risk can be managed and contained in a secure manner, while at the same time ensuring the level of freedom that is necessary for desired scientific advancements. It is our declared intention to raise barriers for malign attackers through a number of measures that will combine to protect Synthetic Biology from abuse. We aim at encouraging continued improvements and harmonization in this field, as well as adoption and further evolution of this Code of Conduct and the Best Practice Guidelines in the future. The Undersigned will participate or otherwise
reasonably contribute for regular scientific dialogue on the further evolution of screening, best practices and the topic of virulence factors and positive or negative lists of elements against which synthetic genes should be screened. The Undersigned promise to develop a compliance plan for adherence to this Code of conduct. This Code has been expressly designed to guide companies and other entities engaged in the synthesis of double stranded DNA of minimum 200 base pairs in length and multi-gene constructs. The Undersigned express no opinion about the extent to which the standards described herein may be applicable to the much shorter sequences known as “oligos.”

3. Risk Assessment and Risk Management

Abuse of synthetic genes in hazardous applications is possible in two ways only: Intentionally, and by failures in risk assessment and management. The technology of handling synthetic genes uses complex procedures which by their nature are self-contained and tightly controlled under existing standards of good practice. For biosecurity, risk assessment entails the screening of DNA sequences for genes which can be intentionally abused, for example, in terrorist activities, whereas risk management entails the restriction of access to synthetic DNA to legitimate users.

4. Record Keeping

- Records of suspicious inquiries and positive screening hits will be kept for at least 8 years.
- Statistics on biosecurity and biosafety related inquiries and orders will be kept for at least 8 years. Information to be retained shall include the total number of inquiries and orders for synthetic genes, the number of inquiries and orders with
positive screening hits, and the number of orders with positive screening hits which have been respectively filled or rejected.

5. **Cooperation with Authorities**

Gene synthesis providers shall take reasonable steps to maintain communications with the government in the nation where they are headquartered. Gene synthesis providers shall promptly inform these authorities each time they encounter evidence which clearly suggests possible illegal activities. Such evidence will include, by way of example, inquiries and orders that strongly suggest illegal activities, such as attempts to conceal a non-business delivery address."

6. **Sequence Screening**

- Gene synthesis companies should always take reasonable steps to determine the relationship of the requested sequences to risk-associated sequences before sending them to customers. The following procedure reflects IASB members’ best collective judgment of how to achieve this goal within the framework of existing technology:
  
  o DNA sequences submitted as inquiries or orders for DNA synthesis by customers will be screened against GenBank for reasonable sequence similarity to pathogens. Members may take further reasonable steps to determine the function and evaluate the associated biorisk associated with homologous genes following procedures to be defined by the Technical Experts Group on Biosecurity (hereinafter “TEGB”). Pending such procedures, providers shall determine and follow their own best practices.
In addition to determining biorisk, entities shall also comply with all national laws. This will include reviewing and comparing top homology hits against (a) all Australia Group biological dual-use organisms, (b) The US Select Agent and Toxins list, and (c) against national organism lists for export control or biological safety/security.

The foregoing procedure establishes a benchmark capability for detecting threat sequences. However we expect researchers to develop new sequence screening technologies over time. Members shall be free to adopt such alternative technologies provided that the new methods have first been empirically shown to detect threat sequences at reliability levels that meet or exceed the benchmark methods described above, as elaborated by TEGB over time. IASB members pledge to promptly update this Code of Conduct to reflect such new (and potentially higher) standards as they appear.

IASB members pledge to take ongoing, collective efforts to refine and improve today’s screening technologies over time. These shall include:

- Establishing a standing Committee to review and if necessary update and extend this Code of Conduct in light of changing threats and/or technology advances over time.
- Regularly exchanging literature searches, virulence judgments, and other data needed to determine the function and/or threat potential of GenBank genes through a secure on-line collaboration to be hosted by the University of California’s Goldman School of Public Policy (VIREP).
• Regularly exchanging, discussing, and collaborating on best practices and ideas through person-to-person contacts and through a secure on-line collaboration.

• Providers which find that a requested gene may code for functions that pose a biosecurity risk shall not fill such orders unless and until they have conducted intensive customer screening at the highest levels provided for in Section 8 of this Code.

7. Response to Identified Threats

• Whenever any of the procedures described in Section 6 produce a “hit” as defined by the then-applicable TEGB guidance, the hit will be assessed by a molecular biologist or similar subject matter expert.

• When the hit is deemed authentic,
  o the customer will be notified and made aware of the perceived risk,
  o the order will be accepted only if the customer is a legitimate user (see section 8) and all national regulations that apply to the exporting/producing company have been met.
  o National authorities shall be contacted as to the extent provided for in Section 5

8. Customer Screening

Gene synthesis providers should always take reasonable steps to confirm that their customers are who they say they are. Where customers seek risk-associated sequences, providers should take further reasonable efforts to confirm that the customer seeks the
requested sequence for legitimate purposes, and has carefully considered any safety or security risks potentially associated with their use of the sequence. The following procedure reflects IASB members’ best collective judgment of how to achieve these goals within the framework of existing technology:

- In a first step, which is to be performed for all orders independent of whether they are considered to be risk-associated:
  - A minimum set of identification data for the customer will be retrieved, including postal address, institution, country, telephone number, and email address.
  - These data will be kept on record according to section 4.
- When an ordered synthetic gene is identified as a risk-associated sequence, the following steps are to be performed:
  - The legitimacy of the customer will be determined by a commercially-reasonable inquiry by the gene synthesis provider and the decision of legitimacy will be documented.
  - It will be ensured that the stated postal address is not a residential address nor a PO box or similar address with limited traceability.
  - The foregoing determination shall include, inter alia, verifying the addresses of businesses and institutions which placed the order, and ensuring that the address owner is a legitimate organization (such as a registered business or an internationally recognized academic institution).
The foregoing procedure establishes a benchmark capability for screening customers. However we expect researchers to develop new screening methods over time. Members shall be free to adopt such alternative methods provided that they meet or exceed the benchmark methods described above. IASB members pledge to promptly update this Code of Conduct to reflect such new (and potentially higher) standards as they appear.

IASB members pledge to take ongoing, collective efforts to refine and improve today’s screening technologies over time. These shall include (a) establishing a standing Committee to review and if necessary update and extend this Code of Conduct in light of changing threats and/or technology advances over time, and (b) regularly exchanging, discussing, and collaborating on best practices and ideas through person-to-person contacts and through a secure on-line collaboration.

IASB members pledge to take ongoing, collective efforts to refine and improve today’s screening technologies over time. These shall include (a) establishing a standing Committee to review and if necessary update and extend this Code of Conduct in light of changing threats and/or technology advances over time, and (b) regularly exchanging, discussing, and collaborating on best practices and ideas through person-to-person contacts and through a secure on-line collaboration. Where the provider’s investigation reveals that its immediate customer of a risk-associated gene is not the intended end-user but will instead re-ship the risk-associated gene to a third party end-user, gene synthesis companies shall either (a) identify and investigate the end-user as provided for in this
Code, or (b) take reasonable steps to confirm that its immediate customer has adopted and routinely follows procedures comparable to those provided for in this Code.

9. **Cooperation on Biosafety and Biosecurity**

- The Undersigned will participate in the formation of a Technical Expert Group on Biosecurity (TEGB). This group will review current design and implementations of biosafety and biosecurity measures, and will propose and initiate improvements.

- The TEGB shall develop an IASB operated seal of approval program to certify compliance with this Code. Providers will be encouraged to apply for seals whether or not they are currently IASB members.
APPENDIX D: IGSC HARMONIZED SCREENING PROTOCOL

Preamble

This document outlines the standards and practices that IGSC gene synthesis companies apply to prevent the misuse of synthetic genes. By screening the sequences of ordered genes and vetting customers, IGSC companies help to ensure that science and industry realize the many benefits of gene synthesis technology while minimizing risk. The ICQS companies together represent approximately 80% of commercial gene synthesis capacity world-wide.

1. Gene Sequence Screening

IGSC companies screen synthetic gene orders to identify regulated pathogen sequences and other potentially dangerous sequences.

1. IGSC companies screen the complete DNA sequence of every synthetic gene order against the DNA sequences in a Regulated Pathogen Database, and against all entries found in one or more of the internationally coordinated sequence reference databanks (i.e., NCBI/GenBank, EBI/EMBL, or DDBJ). The IGSC is currently assembling a Regulated Pathogen Database that will include data from all organisms on the Select Agent list, the Australia Group List, and any other national list of regulated pathogens. Until this is deployed, each company is using its own database of pathogen sequences. At a minimum, IGSC companies screen for all pathogen and toxin genes from
the US Select Agents and Toxins List and/or from the list specified in paragraphs 1C351-1C354 of European Union Council Regulation 428/2009.

2. IGSC companies translate all six reading frames of each synthetic gene into an amino acid sequence. This sequence is screened against the protein sequences derived from the databases described above.

3. IGSC companies use automated screening as a filter to identify pathogen and toxin DNA sequences. When automated screening identifies a potential pathogen or toxin sequence, the order is reviewed by a human expert and is either accepted, accepted with a requirement for additional customer review, or rejected.

2. Gene Customer Screening

1. IGSC companies require identification data from all potential customers for synthetic genes, including at a minimum a shipping address, institution name, country, telephone number, and email address. We do not ship to PO Boxes.

2. Potential customers are screened against OFAC’s SDN List, the Department of State’s Debarred List, and BIS’s Denied Persons, Entity, and Unverified lists, or the HADDEX list, and/or any other list required by applicable national regulations.

3. IGSC companies require additional customer screening before accepting orders for DNA sequences from regulated pathogens. Although the U.S. Select Agent Regulations and the European Commission regulations do not restrict access to all Select Agent genes, IGSC companies supply genes from regulated pathogens only to researchers in government laboratories, universities, nonprofit research institutions, or industrial laboratories demonstrably engaged in legitimate research. Customers ordering Select
Agent or Australia Group DNA fragments must provide a written description of the intended use of the synthetic product; we verify independently a) the identity of the potential customer and purchasing organization, and b) that the described use is consistent with the activities of the purchasing organization.

IGSC companies use the current recommendations from the U.S. CDC and/or the Department of Agriculture and/or the European Commission (CR42) to determine which DNA sequences are Select Agents as recombinant DNA fragments. We supply genes with such sequences only if the supplier and the customer are able to comply with all Select Agent regulations applicable to that gene.

In general, IGSC companies only sell DNA or fragments of regulated pathogens to bone fide end-users. We do not sell or ship such material to distributors or other resellers, unless those companies identify the end-user receiving the products and demonstrate their compliance with every requirement otherwise applicable to that end-user.

3. Record keeping

1. Sequence Screen Results: IGSC companies retain records of every gene sequence screening result for at least 8 years.

2. Customer Screen Results: IGSC companies retain records of every customer screening result for at least 8 years.

3. Product & Delivery Information: IGSC companies retain records of every gene synthesized and delivered for a minimum of 8 years after shipping, including at least the following: (a) the synthetic DNA sequence; (b) the vector; and (c) the recipient’s identity and shipping address.

250
4. Order Refusal & Reporting

1. IGSC companies reserve the right to refuse to fill any order and to notify authorities upon identifying potentially problematic orders.

2. IGSC companies have established relationships with local and national law enforcement and intelligence authorities with whom we can share information to report and to prevent the potential misuse of synthetic genes.

3. IGSC companies will report any request for a gene associated with the pathogenicity of an organism received from a suspicious potential customer and/or potential customer failing to establish its bone fides in application of the practices set forth in section 2.

5. Regulatory Compliance

1. IGSC companies comply with all applicable laws and regulations governing the synthesis, possession, transport, export, and import of gene synthesis and other products.

2. We comply with World Health Organization recommendations concerning the distribution, handling, and synthesis of Variola virus DNA.

Consortium Collaborative Activities

IGSC companies intend to work together in order to:

1. Develop and update a Regulated Pathogen Database to include all gene sequences identified as potentially hazardous by authoritative groups such as the CDC, the Australia Group, and the U.S. and European governments.

2. Ensure that we use the best and most effective algorithms to screen gene sequences against the Regulated Pathogen Database.
3. Collaborate with our respective national governments in support of effective oversight of gene synthesis technology, and to encourage international coordination.

4. Incorporate recommendations from the regulatory, scientific, and public interest communities into our screening and other biosecurity processes.

**Revisions to the Harmonized Screening Protocol**

This document represents an initial effort by a group of companies committed to the responsible use of gene synthesis technology. IGSC companies welcome comments and suggestions to improve the Harmonized Screening Protocol from scientists, regulators, and other interested parties. This document will be revised periodically in response to these suggestions and to changes in the scientific, technical, or regulatory environment.

**Terminology**

Gene Synthesis. This document uses the phrase “gene synthesis” to refer to the production of double-stranded, recombinant DNA fragments from oligonucleotides. Synthetic genes are typically provided in plasmid vectors.

Oligonucleotides. Chemically synthesized, single-stranded DNA fragments, typically supplied as a solution in a tube or a multi-well plate.

Synthetic Gene. A gene or other DNA fragment produced by gene synthesis, typically between 50 and 50,000 base pairs in length.

**Related Links**

Select Agents and Toxins List

EU Council Resolution 428

HADDEX
http://www.ausfuhrkontrolle.info/ausfuhrkontrolle/de/arbeitshilfen/haddex/index.html

OFAC’s SDN List
http://www.treas.gov/offices/enforcement/ofac/sdn/

Department of State’s Debarred List
http://www.pmddtc.state.gov/compliance/debar.html

BIS’s Denied Persons, Entity, and Unverified lists
http://www.bis.doc.gov/complianceandenforcement/liststocheck.htm

Current Recommendations from the U.S. CDC
http://www.selectagents.gov/SyntheticGenomics.html

Australia Group Listed Source Organisms
http://www.australiagroup.net/en/biological_agents.html

World Health Organization Recommendations Concerning the Distribution, Handling, and Synthesis of Variola Virus DNA
http://www.who.int/csr/disease/smallpox/SummaryrecommendationsMay08.pdf
APPENDIX E: INTERVIEW QUESTIONS FOR GENE SYNTHESIS FIRMS

1. How would you describe your work?
2. How many hours per week do you spend conducting laboratory work?
3. How many years of experience do you have in the bench laboratory sciences?
4. What is the highest level of education you have attained?
5. What types of techniques/procedures do you use in your work?
6. How many of these assays are new/not well established? How many are well established?
7. How often do you use these techniques/methods?
8. Which techniques/method require the use of instrumentation?
9. Please describe the amount of human involvement required to use laboratory instrumentation.
10. How often do you rely on trial and error to resolve technical difficulties?
11. How often do you rely on protocols?
12. How often do you rely on manufacturer instruction manuals when using laboratory equipment?
13. Do manufacturer instruction manuals for laboratory equipment ever need to be modified or augmented?
14. Generally speaking, does following protocols typically produce reliable results, or do protocols need to be modified in order to obtain the desired results?

15. Do protocols need to be modified in order to obtain the desired results? If so, how often?

16. If protocols need to be modified, where does the information come from to modify them?

17. By what means do new and/or inexperienced personnel learn how to conduct the techniques you routinely conduct?

18. Does the improper use of laboratory instrumentation ever affect laboratory results? If yes, please provide a specific example.

19. How long does it take for new and/or inexperienced personnel to learn how to perform laboratory tasks?

20. How often do contingencies arise when performing tasks?

21. How do you resolve unexpected problems in performing laboratory tasks?

22. How often do unexpected problems arise when using automated machines in the laboratory?

23. How are these unexpected problems resolved?

24. How compartmentalized is your company? Are you encouraged to network with employees in other divisions, employees of similar companies, or other professionals?

25. How would you characterize your management style?
REFERENCES


Smith, H.O., C. A. Hutchinson, C. Pfannkoeh, and Craig Venter. “Generating a Synthetic Genome by Whole Genome Assembly: pXi74 Bacteriophage from Synthetic


270


Shannon R. Fye-Marnien graduated from California State University, California in 2004. She received her Masters of Science from George Mason University in 2006. She currently works as an analyst for The Tauri Group in Alexandria, Virginia, where she conducts studies on technologies for government and commercial clients.