Data Brief

# Mapping Biosafety Level-3 Laboratories by Publications

**Authors**
Caroline Schuerger
Sara Abdulla
Anna Puglisi

CSET CENTER for SECURITY and EMERGING TECHNOLOGY

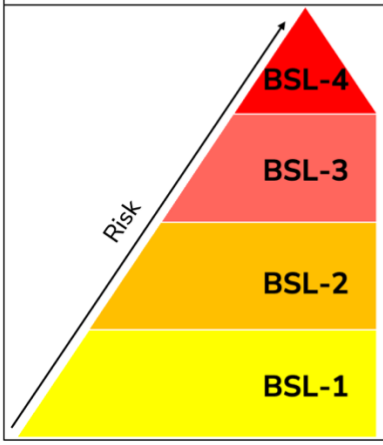
Aug 2022

## Introduction

Understanding and characterizing disease-causing organisms present a distinct dilemma for the scientific and policy communities. On the one hand, this research presents great opportunities for developing vaccines and therapies, understanding host-pathogen interactions, and possibly preventing future pandemics. On the other, it potentially opens the door to lab accidents or potential misuses, just the fear of which may spur governments and regulatory bodies to enact over-restrictive rules and regulations that thwart discovery rather than protect society. A key part of this research infrastructure are Biosafety Level-3 laboratories, which enable scientists to research high-containment infectious pathogens in a safe and secure facility. Therefore, understanding the number and distribution of BSL-3 laboratories provides insight into where this research is being conducted.

Despite their importance, there is currently no comprehensive list of BSL-3 facilities, or the institutions in which they are housed. As a way to begin to better understand the depth and breadth of this research, as well as how it is distributed worldwide, we systematically identified and located BSL-3 institutions that published in English from 2006–2021, using PubMed Central to identify those publications. This paper maps institutions that host a BSL-3 laboratory by location to provide a high-level analysis of where high-containment BSL-3 research is conducted globally.

Figure 1: Biosafety Levels

| Biosafety Level (BSL) | Characteristics | Pathogens/Disease |
|---|---|---|
|  | Infectious aerosol transmission that may cause serious or lethal infections with no treatment available | Ebola virus, Variola virus (smallpox), Marburg virus |
| | Infectious aerosol transmission that may cause serious or lethal infections | Coronavirus, Mycobacterium tuberculosis, Yersinia pestis (plague), malaria |
| | Infectious agents of moderate risk with ingestion or mucous membrane transmission | Influenza, Lyme disease, salmonella, measles, mumps |
| | Low-risk agents that are not known to cause human disease | E. coli |

Source: Adapted from the United States Center for Disease Control and Prevention.

## Current Biosafety Policies

Current U.S. policies, programs, and regulations do not capture the full extent of domestic BSL-3 research because they are responsible for high-containment oversight in specific situations. The first is the Select Agent Program, which requires registration if a laboratory is conducting research on a pathogen designated as a select agent. This list includes pathogens that require a range of biosafety levels, but some pathogens that require BSL-3 research precautions are not included.[1] The second is the National Institute of Health's Biosafety and Recombinant DNA Policy, which oversees institutions that receive federal funding for experiments that involve recombinant DNA, including BSL-3 work.[2] Since many research experiments with BSL-3 pathogens fall outside of these two programs, some laboratories researching BSL-3 pathogens are unregistered with the U.S. government.[3]

Globally, while there are World Health Organization guidelines and biosafety best practices, there is no international body that monitors or regulates high-containment lab research. The WHO publishes best practices and guidance for biocontainment, but it is up to the country and researcher to implement proper biosafety, further obscuring BSL-3 research.[4]

## Non-Publishing Institutions: Reference and Clinical Diagnostic Laboratories

Many BSL-3 institutions are National Reference Laboratories and clinical diagnostic facilities established through domestic and international collaborations, and provide diagnosis and reference information for diseases. These BSL-3 institutions usually function outside of academic settings and are less likely to publish, and will not be reflected in our analysis of publications available through PubMed Central. For example, the U.S. Agency for International Development supports efforts in many countries to establish, develop, and maintain the safety of BSL-3 reference laboratories. In Kyrgyzstan and Zambia, they were established to diagnose and surveille endemic tuberculosis, and provide training and resources for laboratory practitioners.[5] Additionally, the U.S. Defense Threat Reduction Agency of the Department of Defense also works with different countries to establish BSL-2 and -3 laboratories as part of the Central Public Health Reference Laboratory. These include facilities in the Republic of Georgia and Kazakhstan.[6] U.S. cooperation and outreach to establish health-related BSL-3 institutions provide clinical resources and suitable biosafety resources as a way to build capacity while preventing lab accidents.

International institutions also build and administer reference and diagnostic labs. In Africa, the World Bank funded BSL-3 reference laboratories, supporting the Africa Centres for Disease Control and Prevention through the Africa CDC Regional Investment Financing Project.[7] In 2019, this program launched regional reference laboratories in Ethiopia and Zambia, specifically to work with tuberculosis, HIV/AIDS, and malaria. The Africa Pathogen Genomics Initiative is building BSL-3 scientific capacity to support disease diagnosis and public health surveillance instead of basic research, and therefore these laboratories are unlikely to have publications in our data.[*8] These Reference Laboratories and clinical diagnostic centers are part of the larger landscape of BSL-3 institutions to be identified in future work.

## Methodology

To identify published research performed under BSL-3 containment, we performed a keyword and pattern search over a dataset of about two million English-language articles published between 2006 and 2021, which we retrieved from the PubMed Central Open Access Subset and Author Manuscript datasets.[†] The data included full-text content along with other metadata.[‡]

Among the publications that mentioned BSL-3 containment, we looked in the text for an indication of which institution's laboratory performed the work. This step allowed us to positively identify institutions with BSL-3 facilities, relying on self-reporting by authors. We inspected each of these results for accuracy by examining each publication's methodology for utilizing a BSL-3 laboratory. Our approach emphasized precision over recall, in the sense that we will not have identified all the references to BSL-3 facilities in the PubMed Central data, but we have high confidence in each

[*] The Africa Pathogen Genomics Initiative was launched by the Africa CDC. Contributions to the Africa Pathogen Genomics Initiative include funding and technical assistance from the U.S. CDC and the Bill & Melinda Gates Foundation, and technical assistance and resources for next-generation sequencing machines or data architecture from Illumina, Oxford Nanopore, and Microsoft.

[†] We searched for several possible variations of "biosafety level 3" and "BSL-3" after a review of the literature that identified references in common usage. For example, some authors refer to "biologic" or "biological safety level" rather than "biosafety level." We saw "BSL-3" with or without a hyphen. Our implementation of this flexibility used regular-expression matching against two patterns: 'bio(logic|logical)*[\s-]+safety[\s-]+level' and 'BSL[\s-]*3'. We searched against the following fields in PubMed data, ignoring case: 'full_title', 'vernacular_title', 'abstract', 'other_abstract', and 'text'.

[‡] For background, see "About PMC," PubMed Central, U.S. National Library of Medicine, March 21, 2022, https://www.ncbi.nlm.nih.gov/pmc/about/intro/ and "Accessing PMC Article Datasets Using Amazon Web Services," PubMed Central, U.S. National Library of Medicine, June 17, 2022, https://www.ncbi.nlm.nih.gov/pmc/tools/pmcaws/.

identification made. Our final step was to identify the locations in which these institutions were located.

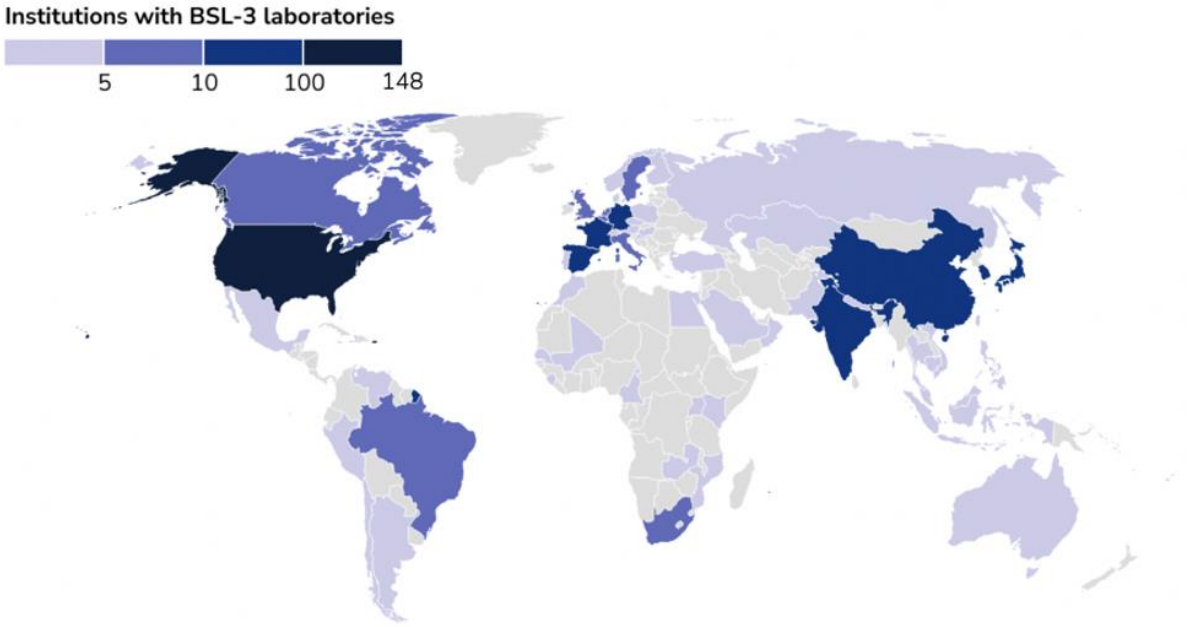
Prior research has relied primarily on surveying or investigating known institutions to locate and characterize BSL-3 facilities, resulting in ad hoc collection of high-containment institutions.[9] By applying text mining methods to the scientific literature, we have identified laboratories, both in the United States and abroad, that add to the prior research to include institutions that publish BSL-3 research that might not have been captured in previous methods. This is only one approach of identifying BSL-3 institutions, and we plan to use additional methods to further characterize global institutions.

This analysis does not include BSL-3 laboratories that do not publish their research. It also excludes BSL-3 laboratories whose publications do not appear in PubMed Central despite whether their research is published or not, as well as those whose publications appear in PubMed Central but are not written in English, and laboratories that study BSL-3 pathogens that do not self-report using BSL-3 biocontainment methods in the methodology section of their publication. On the other hand, our approach is able to identify BSL-3 facilities that are less well-known but whose research has implications for biosafety and biosecurity just the same.

## BSL-3 Laboratory Locations

The majority of BSL-3 institutions we identified through their publications are concentrated in the United States and China, with Europe representing the next tier.* Of the 58 locations represented in our findings, only eight have more than 10 unique institutions with BSL-3 laboratories.

Figure 2: Location of BSL-3 Institutions Worldwide




Source: PubMed Central, CSET analysis.

There are many factors that contribute to the high number of U.S. BSL-3 institutions that publish. We assess that the major factor for the high level of BSL-3 publishing institutions in the United States as identified by our methodology is the federal oversight in the U.S. research life cycle by federal grant policy requirements. Federal grant policies require that researchers not only publish their findings but often require their data, including detailed descriptions of experimental methods, to be posted as well.†

* See Daniel Chou, "Counting AI Research: Exploring AI Research Output in English- and Chinese-Language Sources" (Center for Security and Emerging Technology, July 2022), https://cset.georgetown.edu/publication/counting-ai-research/. We currently have a follow-on study looking at specific institutions that publish in Chinese.

† The National Institutes of Health Public Access Policy and the Office of Science and Technology Policy Public Access Memorandum requires federally-funded research be open-access.

Our data also finds that in the United States there is a broader representation among the types of BSL-3 publishing institutions, perhaps reflecting the decentralized nature of the U.S. research and development (R&D) system. The 148 BSL-3 institutions we found in this study include federal research centers, universities, and companies. They are also distributed across the United States geographically.

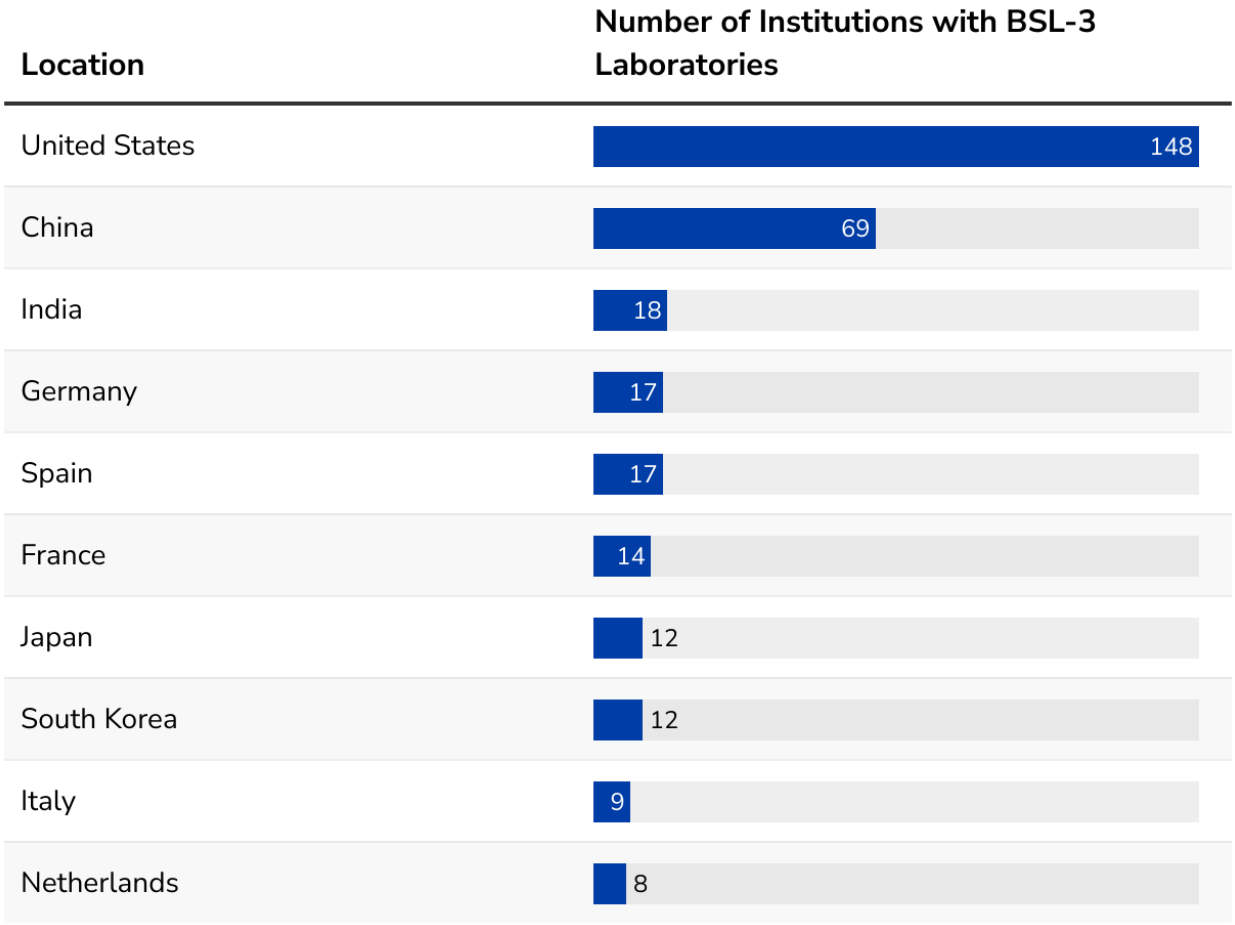
China's prioritization of biotechnology R&D may contribute to the country having the second largest number of BSL-3 institutions in our analysis. The Chinese government directs research, including biotechnology, to meet its strategic goals and address the economic and societal needs in China. Policies such as the Medium- and Long-Term Plan for Science and Technology (S&T) Development (2006-2020) and the 13th Five-Year Plan for S&T Innovation directly outline steps to develop biotechnology, including exponentially increasing R&D funding investments and establishing innovation hubs.[10] Specifically, infectious diseases and public health are among the central government's research priorities. The S&T plans indicate that China continues to build laboratory capacity to meet its strategic goals, and invest in major programs, which could be why they have so many labs.

In contrast to the United States, our data shows the types of BSL-3 institutions in China are homogenous, reflecting a centralized system of R&D. These institutions consist mainly of state-controlled institutions, such as State Key Laboratories and People's Liberation Army military hospitals, as well as research universities that have close ties to the central government.*

After the United States and China, there is a steep drop-off in BSL-3 publishing institutions elsewhere. India, Germany, Spain, France, Japan, South Korea, Italy, and the Netherlands all have between 8 and 18 BSL-3 publishing institutions (Figure 3). Twenty-five locations have between two and eight institutions, and 23 locations more have only one institution represented in the data (see appendix).

* These universities are almost all "double first class" universities as part of China's education initiatives and are coauthors on BSL-3 publications with State Key Laboratories and PLA Military Hospitals.[11]

Figure 3: BSL-3 Labs by Location




Source: PubMed Central, CSET analysis.

We identified only four institutions in Russia that published BSL-3 work in English. All of these institutions are federal research centers. This may reflect a different threshold for publishing in this field, a preference for publishing in the Russian language or an emphasis on research that is either classified or for military purposes. Generally, Russia publishes fewer papers in English compared to other countries such as China and India. This could also reflect the current state of Russian science where restructuring of scientific institutions, combined with a decline in the number of scientists could contribute to fewer institutions with BSL-3 laboratories.[12]

### *Representation by Region*

Individual European countries have few BSL-3 institutions, but together they represent about a quarter of all institutions globally. Germany, Spain, France, Italy, and the Netherlands have over eight BSL-3 institutions, with the remaining 13 locations represented in the data having at least one BSL-3 institution (Figure 4).

Figure 4: BSL-3 Laboratories by Continent

| Region | Number of Institutions with BSL-3 Laboratories | Number of Locations |
|---|---|---|
| North America | 157 | 4 |
| Asia | 140 | 18 |
| Europe | 102 | 18 |
| Africa | 20 | 11 |
| South America | 12 | 5 |
| Oceania | 3 | 1 |

Source: PubMed Central, CSET analysis.

Though only four locations in North America are represented in the data, more institutions there publish BSL-3 laboratory work than in any other continent. Of 157 BSL-3 institutions in North America, 148 institutions are located in the United States, while the remaining are located in Canada, Mexico, and Haiti. The BSL-3 research and diagnostic institution in Haiti, GHESKIO, is affiliated with Weill Cornell Medicine in New York.

Asia has 18 different locations in which we identified BSL-3 institutions, with China making up the bulk of the total count. Africa has 11 countries represented, with South Africa containing the most BSL-3 institutions in the region with a total of seven. South America has five countries represented in our analysis, with Brazil contributing five BSL-3 institutions. Australia has three separate publishing institutions, and is the only location in Oceania found in our analysis.

In analyzing BSL-3 institutions by regions, we see that one location makes up the bulk of institutions in North America and Asia, but not in Europe, Africa, or South America. BSL-3 research at large institutions in Europe and Africa could serve as a regional space for nearby countries to do their BSL-3 research. Erasmus Medical Center in the Netherlands collaborates with several European institutions, as well as the WHO Tuberculosis Reference Laboratory in Uganda. Researchers in these regions could therefore be less likely to have their own BSL-3 facilities as they rely on these larger regional institutions. Interestingly, the United States has several "Regional Biosafety Centers," where researchers can conduct BSL-3 research in a high-containment setting, but the United States still has a large total institution count.

## Conclusion

The BSL-3 institution locations we found by analyzing English-language publications give us more insight into research on pathogens and research infrastructures around the world, adding to previous investigations. Previous BSL-3 institution discovery located individual laboratories in the United States, rather than counting institutions as in our study. The Government Accountability Office counted over 1,300 BSL-3 laboratories registered with the Select Agents Program, and a *USA Today* investigation found over two hundred total BSL-3 and -4 laboratories around the United States. These studies count individual laboratories instead of institutions, but these counts can be unclear if multiple laboratories use the same biosafety facility, or if the said laboratory moves to a different institution. Additionally, laboratories that work with select agents might not publish, and exclude other BLS-3 pathogen research.

The U.S. BSL-3 institutions consist of a mixture of federal research centers, universities, and companies, while China's BSL-3 institutions are mostly state-level institutions, including State Key Laboratories, military hospitals, and collaborating universities. These differences in types of BSL-3 institutions in our data probably reflect the different approaches to research in the two countries.

Our study also indicates that other regions have fewer numbers of BSL-3 institutions but frequent partnerships between institutions in different locations. The differences in BSL-3 institution counts in individual countries in Europe compared to the whole European region points to frequent collaborations. Places where many BSL-3 pathogens are endemic, such as areas in Africa and Central Asia, do not have many BSL-3 institutions that publish. There, most BSL-3 resources might be diverted to reference laboratories or clinical diagnostics centers.

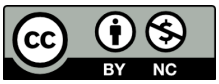
This brief investigated BSL-3 institutions using English-language publications to start chipping away at locating who and where BSL-3 research is being conducted. Further work will identify institutions that conduct BSL-3 work but do not publish in English in PubMed, whether it is due to other institutional functions such as clinical diagnostics or publications in non-English languages. Together, a full picture of BSL-3 research will aid policymakers and governing bodies to make informed decisions on facilitating and regulating high-containment research.

## Authors

Caroline Schuerger is a biotechnology research fellow at CSET, where Sara Abdulla is a data research analyst, and where Anna Puglisi is a senior fellow and the director of biotechnology programs.

## Acknowledgments

The authors are especially grateful to Alison Young and Norm Kahn for their comprehensive reviews, and James Dunham for data support and ideas. The authors would like to thank Catherine Aiken and Igor Mikolic-Torreira for their feedback and comments on this report. Thanks also to Lynne Weil, Shelton Fitch, and Jess Shao for their assistance in reviewing, editing, and formatting this paper.





Document Identifier: doi: 10.51593/20220019

## Appendix

Table 1: Location Count of Publishing BSL-3 Institutions

| Location | Institutions |
|---|---|
| United States | 148 |
| China | 69 |
| India | 18 |
| Germany | 17 |
| Spain | 17 |
| France | 14 |
| Japan | 12 |
| South Korea | 12 |
| Italy | 9 |
| Netherlands | 8 |
| South Africa | 7 |
| Sweden | 7 |
| Canada | 7 |
| Belgium | 6 |
| United Kingdom | 5 |
| Brazil | 5 |
| Indonesia | 4 |
| Taiwan | 4 |
| Malaysia | 3 |
| Turkey | 3 |
| Denmark | 3 |
| Hungary | 3 |

| | |
|---|---|
| Switzerland | 3 |
| Australia | 3 |
| Argentina | 3 |
| Kenya | 2 |
| Uganda | 2 |
| Zambia | 2 |
| Singapore | 2 |
| Thailand | 2 |
| Austria | 2 |
| Croatia | 2 |
| Finland | 2 |
| Peru | 2 |
| * | 1 |

* The locations with one BSL-3 publishing institution were: Cameroon, Cambodia, Chile, Czech Republic, Egypt, Haiti, Kazakhstan, Mali, Mexico, Morocco, Mozambique, Nepal, Norway, Oman, Pakistan, Philippines, Poland, Portugal, Saudi Arabia, Senegal, Sierra Leone, Venezuela, Vietnam.

Source: PubMed Central, CSET analysis.

## Endnotes

[1] "Select Agents and Toxins List | Federal Select Agent Program," March 30, 2022, https://www.selectagents.gov/sat/list.htm.

[2] "NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules (NIH Guidelines) - April 2019," 2019. https://osp.od.nih.gov/wp-content/uploads/NIH_Guidelines.pdf

[3] U.S. Government Accountability Office, "High-Containment Laboratories: Coordinated Actions Needed to Enhance the Select Agent Program's Oversight of Hazardous Pathogens," accessed May 3, 2022, https://www.gao.gov/products/gao-18-145.

[4] World Health Organization, ed., *Laboratory Biosafety Manual*, 4th ed (Geneva: World Health Organization, 2020).

[5] USAID, "TB Labs in Kyrgyzstan Achieve International Standards | Transforming Lives | Kyrgyz Republic | U.S. Agency for International Development," March 23, 2017, https://www.usaid.gov/results-data/success-stories/tb-laboratories-kyrgyzstan-meet-international-quality-standards; USAID, "Zambia Partnership Statement," January 16, 2020, https://www.usaid.gov/global-health/health-areas/tuberculosis/zambia-partnership-statement.

[6] Department of Defense, "The Birth of a Laboratory," U.S. Army, 2012, https://www.army.mil/article/82553/the_birth_of_a_laboratory; "BSL 3 Research Laboratory Opens in Kazakhstan," *Global Biodefense*, August 27, 2014, https://globalbiodefense.com/2014/08/27/bsl-3-research-laboratory-opens-kazakhstan/.

[7] The World Bank, "Africa CDC Regional Investment Financing Project," https://projects.worldbank.org/en/projects-operations/project-detail/P167916.

[8] Africa CDC, "US$100 Million Africa Pathogen Genomics Initiative to Boost Disease Surveillance and Emergency Response Capacity in Africa," *Africa CDC* (blog), 2020, https://africacdc.org/news-item/us100-million-africa-pathogen-genomics-initiative-to-boost-disease-surveillance-and-emergency-response-capacity-in-africa/.

[9] U.S. Government Accountability Office, "High-Containment Laboratories: National Strategy for Oversight Is Needed," accessed June 9, 2022, https://www.gao.gov/products/gao-09-574; Alison Young and Nick Penzenstadler, "Biolabs in Your Backyard," USA Today, 2015, https://www.usatoday.com/pages/interactives/biolabs/; Filippa Lentzos and Gregory D. Koblentz, "Mapping Maximum Biological Containment Labs Globally," 2021; National Academies of Sciences, *Developing Norms for the Provision of Biological Laboratories in Low-Resource Contexts: Proceedings of a Workshop*, 2019, https://doi.org/10.17226/25311.

[10] Original CSET Translation of "National 13th Five-Year Plan for the Development of Strategic Emerging Industries," [国务院关于印发“十三五”国家战略性新兴产业发展规划的通知], Central People's Government of the People's Republic of China, November 29, 2016, https://cset.georgetown.edu/research/national-13th-five-year-plan-for-the-development-of-strategic-emerging-industries/; Original CSET Translation of "2019 Nationwide Statistical Communiqué on Science and Technology Spending and Investment" [2019年全国科技经费投入统计公报], National

Bureau of Statistics Website, August 27, 2020, https://cset.georgetown.edu/wp-content/uploads/t0229_2019_statistical_communique_EN.pdf.

[11] Ryan Fedasiuk, Alan Omar Loera Martinez, and Anna Puglisi, "Sustaining U.S. Competitiveness in Semiconductor Manufacturing" (Center for Security and Emerging Technology, March 2022), https://doi.org/10.51593/20210007.

[12] Raymond A. Zilinskas and Philippe Mauger, "Biosecurity in Putin's Russia" (Lynne Rienner Publishers, Inc., 2018).