Towards a Grand Unified Threat Model of Biotechnology

By: Michael Montague, PhD
Michael.Montague@jhu.edu
Michael_G_Montague@yahoo.com
Senior Scholar, Center for Health Security, Johns Hopkins
Research Analyst, Future of Humanity Institute, Oxford

Abstract

In the absence of empirical data concerning the capabilities of modern biotechnological methods to produce and deploy high impact biological threat agents, a strong theoretical model is required to inform effective biotechnological regulations and biosecurity preparations. Such a model is presented that aims to be robust across the diverse natures of all biological agents, any actors who might develop them, and the many biotechnologies and emerging computational super intelligence platforms that might be harnessed to do so. Core to this model is the recognition that any high consequence biotechnological agent must be able to spread geographically, be novel to the defenders, and be produced within the well understood constraints of technological development pipelines. Given these requirements, and the well established difficulty of modeling, and manipulating a novel organism's dynamics when introduced into an ecosystem, it becomes possible to derive the necessary properties of any actor capable of developing such a high consequence biotechnological threat agent: They must be designing their agent deliberately to do harm, and they must be highly resourced. Malevolent low resourced actors and benevolent or accidental actors regardless of resource level are revealed as being unable to produce such an agent. This is significant as much recent concern over the democratization of biotechnological capabilities has focused upon the large numbers of potential actors in those categories. Additionally, the constrained nature of the research and development efforts that might actually be able to produce a high consequence biotechnological threat agent allows for a refined focus in biosecurity policy and biotechnology regulation. This refined focus de-emphasizes damaging access-control policies seeking to limit and control large numbers of actors in the biotech space. Instead, an emphasis upon intelligence gathering to detect the definable and large footprints of the kind of research and development program needed to create such a high consequence biotechnological threat agent is revealed as optimal.

Introduction

In 2020, Sandberg and Nelson[1] argued that the democratization of biotechnology has multiplied the number of potential malicious or incompetent actors such that, although any individual disgruntled graduate student or citizen biohacker might represent a very small threat, as a class, they represent the bulk of the danger of misuse of biotechnology. Their 2020 analysis was based upon a "biorisk chain" that describes the advancement of a biological agent's development from concept to deployment to consequences. Importantly, they noted in discussing their conclusion that if one step of the biorisk chain turned out to be much more difficult than any of the others it might invalidate their conclusion as to the nature of the actors of greatest threat. Here an analysis of the properties of high consequence biotechnological agents argues that there *is* a single step in the biorisk chain that is in fact massively more difficult than all others, and that it *does* significantly alter the expectation of what classes of actors represent the majority of the biotechnological threat.

This analysis aims to be independent of the properties of any individual biological agent. Rather, it is based upon an observation of the properties that set all biological agents apart from chemical, kinetic, or radiological agents, as well as properties universal to all technological development cycles. These observations are then contrasted against the capabilities of various classes of actor.

The Necessary Properties of a High Consequence Biotechnological Threat Agent

In order for a biological threat agent that is developed by human actors through a technological process to have the potential for very high consequences, in the range refereed to as Global Catastrophic Biological Risks[2] or Existential Risks[3] it must have all three of the following:

1. The agent must have the capacity to spread. The term "spread" is carefully chosen rather than other terms such as "contagion", as it does not specify whether spread is across patients, species, genetic diversity, demographics, etc. Regardless of the underlying mechanism of a biological agent's dissemination, for it to be the sort of high consequence spread discussed here it must amount to geographic spread across large areas. For example, an invasive species and a pathogen both spread geographically, but by completely different mechanisms.
2. The agent must be novel. Novelty is not a binary either-or property but rather exists on a spectrum. For example, a vaccine evasive variant of a known pathogen is somewhat novel in that it requires some novel responses (a new vaccine). But it likely does not require *entirely* new responses. Lessons learned from prior variants about case mortality rates, patient risk factors, modes of transmission, standards of personal protective equipment, etc might still be valid to some degree making the variant less novel than a completely unprecedented pathogen would be. The degree of threat that an agent represents is, all other things being equal, directly proportional to the degree that it is novel. This is because each unknown concerning the agent's properties represents a barrier discovery before and effective response can be marshaled.
3. The biotechnological agent must be developed, like every other technological product, through the constraints of the iterative Design > Build > Test > Learn technological development cycle.

These three properties[1] of a high consequence biotechnological agent, and the required capabilities of a development pipeline that could produce it, circumscribe and restrict the properties of the kind of actor that could build and use such a pipeline.

Spread and it's Implications to Biotechnology Threats

Spread is the definitive property that many biological agents can posses, and that no non-biological agents do. Still, it should be noted that not all biological agents need to be confined to a spread-based usage mode, and in fact many known bioweapon agents have historically been deliberately chosen or engineering to achieve a *lack of spread* because that makes the agent a more controllable weapon; anthrax spores are one example of this.[4] However in such cases, that lack of spread means that the agent is only as powerful as its delivery mechanism, or no different from kinetic, nuclear, or chemical, agents. Thus, if such an combined delivery and payload system is of high consequence, it is not a function of the agent-payload, but only because the delivery system is able to apply it either broadly or specifically.

Rather, the sort of high consequence biological agent that is feared to have global effects represent agents that can, without special delivery mechanisms, reach far beyond the parameters of their initial deployment. That is, they spread themselves. Biological systems have the potential to achieve this through the mechanism of self-replication. But, <u>replication, and thus spread, is never a stand-alone feature</u>. This is true almost regardless of the exact nature of the agent or its replication strategy[2]:

- A pathogen can not replicate except in a compatible host.
- A prion can not replicate except in a tissue that supplies the non-aberrant form of the prion.
- A gene drive can not replicate itself into the next generation of a host organism, and thus into an ever larger fraction of the host's gene-pool, unless it leverages the replication of that host-organism.
- An invasive species can not replicate itself absent a permissive non-competitive ecological niche.

In looking at these very diverse examples, we can see the unifying property that not only is replication, and thus spread, not a stand-alone feature, it is dependent upon far more complexity in the environment that is to be replicated in, than there are features or complexity in the agent itself. As an example,

---

1 Note that the lethality or specific consequences of the agent is not defined on the list of required properties of a high impact biotechnological agent although it is presumed that the agent will have some sort of consequence. This is mostly because the exact nature of the impact is immaterial to the reasoning of the threat model. However, measurements of lethality is often incorrectly privileged as an equal or even more important factor in considering biological agents. The reader is invited to consider which has killed more people: COVID with a case mortality rate well below 1% but with the ability to spread human-to-human around the globe in mere months, or rabies with a case mortality rate approaching 100%, and dependent upon rare inefficient animal vectors for human infection and thus almost no global spread potential?

2 There *are* a very few organisms such as algae, lichen, and certain chemotrophs that represent the base of the food chain and thus could be said to have relatively stand-alone replication capabilities. They do not generally get considered as potential bases of biothreat agents however, and for a very good reason: In order for them to be truly independent of support from the wider ecosystem, they are forced to genetically encode *all* of the genes needed to support all of their life-functions in all environments they might be able to invade. Thus they are either extremophiles that occupy narrow ecological niches that afford them freedom from competition at the expense of not being competent to spread outside of that niche and/or they have slow metabolic life cycles as a consequence of having to carry the metabolic load of all of that genetic capability for self-sufficiency in a wide variety of ecological niches. Either way, they are relatively incompetent at spreading through an environment compared to an organism that does not try to do everything itself which is exactly what any hypothetical bio-agent based upon a stand-alone-replication competent organism would have to compete against once outside the lab or extreme environment.

SARS-CoV-2 has 29.9 thousand bases of genome that enclose 25 genes (counting each of the poly protein digest products of Orf1 as separate genes).[5] The human organism that it replicates in, by contrast, has a genome of 2.91-billion base pairs and on the order of 27000 genes.[6] That is 3-6 orders of magnitude more complexity. Similarly, the prion relies upon a tissue that can transcribe and translate more copies of the non-aberrant form; a series of processes orders of magnitude more complex than the auto-catalysis of the prion's own replication process in that environment. A gene drive, properly speaking, doesn't replicate itself so much as stow-away on or trigger the replication and/or recombination machinery of the host organism which is itself much more complex than the gene drive. Similarly the ecological niche that an invasive organism dominates is maintained by the life-processes of thousands or even millions of other organisms in that ecosystem.

This general observation allows us to make a conclusion about any as yet to be observed future biological agents whether they emerge naturally or are developed through a technological process: The dynamics of how or whether any future biological agent spreads in an environment will be at least 99% a function of the complexity and properties of that environment with only a small fraction of the agent's spread potential being a direct and exclusive application of the properties of the agent itself.

This conclusion has profound implications for anyone trying to create such an spreading agent through a biotechnological process. First, it means that the spread dynamics of an agent simply can not be calculated from the design of the agent itself. Consider once again the case of SARS-CoV-2. More than any other specific trait that it possesses, the ability to be contagious without visible/severe symptoms is arguably the one most responsible for making SARS-CoV-2 a pandemic-capable pathogen.[7]–[9] Which of its 25 protein genes grants it that trait? Insofar as there can be an answer, it is all or most of them, or perhaps none of them; after all, symptom-less spread in humans is at least partly, probably mostly, a function of some of the ~27,000 human genes[10], [11], and of course partly encoded by non-genetic behavioral traits like masking[12], [13]. Consequently, no understanding of the 25 SARS-CoV-2 genes in isolation, no matter how complete or sophisticated, could possibly recapitulate, or likely even guess at that pandemic enabling spread trait.[3]

Second, because these environments that a biological agent would spread in are almost always insufficiently sampled to characterize in real time, or possibly at all, developing an agent that can spread in them, to say nothing of spread with characteristics engineered to be predictable and desirable is not something that can be done based upon calculation from first principles no matter how well understood the agent's own biology is. The only way to design or predict such spread characteristics of a novel agent in the wild is to *test* and tune its behavior either in the actual system it will spread in, or in a very near proxy.

It is important to make a distinction here: the testing referred to here is not viability testing which merely shows that the agent can function and survive in the controlled conditions supplied by a laboratory without immune responses, predators, competitors, or confounding factors. Rather, the focus here is on ***spread testing*** to determine the if and how a engineered bioagent will or will not spread upon deployment in the real world when confronted with all those confounding properties.

Unlike viability testing, testing for spread dynamics is and must be complex, slow, hard, and expensive. This is because the spread of an agent in a wild real world setting is dominated by all of the unmanageable large-scale outside factors that laboratories are specifically designed to exclude so that they can perform controlled, affordable, reproducible experiments. Consider the difficulty of working around this in the example of a hypothetical engineered human pathogen: At a minimum, spread-testing such a novel engineered human pathogen would require some way to duplicate or recapitulate the

3 This is analogous to the math problem Y=X+7. No amount of mathematical knowledge can make this a solvable equation in isolation. The necessary information for the solution simply is not present in the provided problem. Even a super-intelligent AI of arbitrary capability could not solve such data-limited problems.

dynamics of human to human spread; that requires an animal model at the very least. But finding and validating an animal model is extremely hard, and even when it exists, the model is rarely as predictive as one would like.[14] Further, humans and animals in the wild do not behave as animals or humans in cages behave. The result is that while animal models provide a mediocre proxy for organism viability, their capacity to model organism to organism spread is much less predictive. Again SARS-CoV-2 demonstrates the point: Animal models can be created that recapitulate some of the dynamics of infection and replication, but they all differ in disease severity or presentation[15] (both of which are factors affect how the pathogen spreads in humans). The spread of biological systems like gene drives or invasive species are notoriously hard to model. [16]–[20] All of this falls under the general field of managing managing ecosystems which is widely recognized as a "wicked problem"[21] that is not amenable to the kinds of the kinds of simple rules and predictions associated with more tractable sciences[22]. An interesting exception to the outside world be characterized by environments that are refractory to laboratory spread testing is agriculture; this will be discussed further in the discussion section.

Figure 1

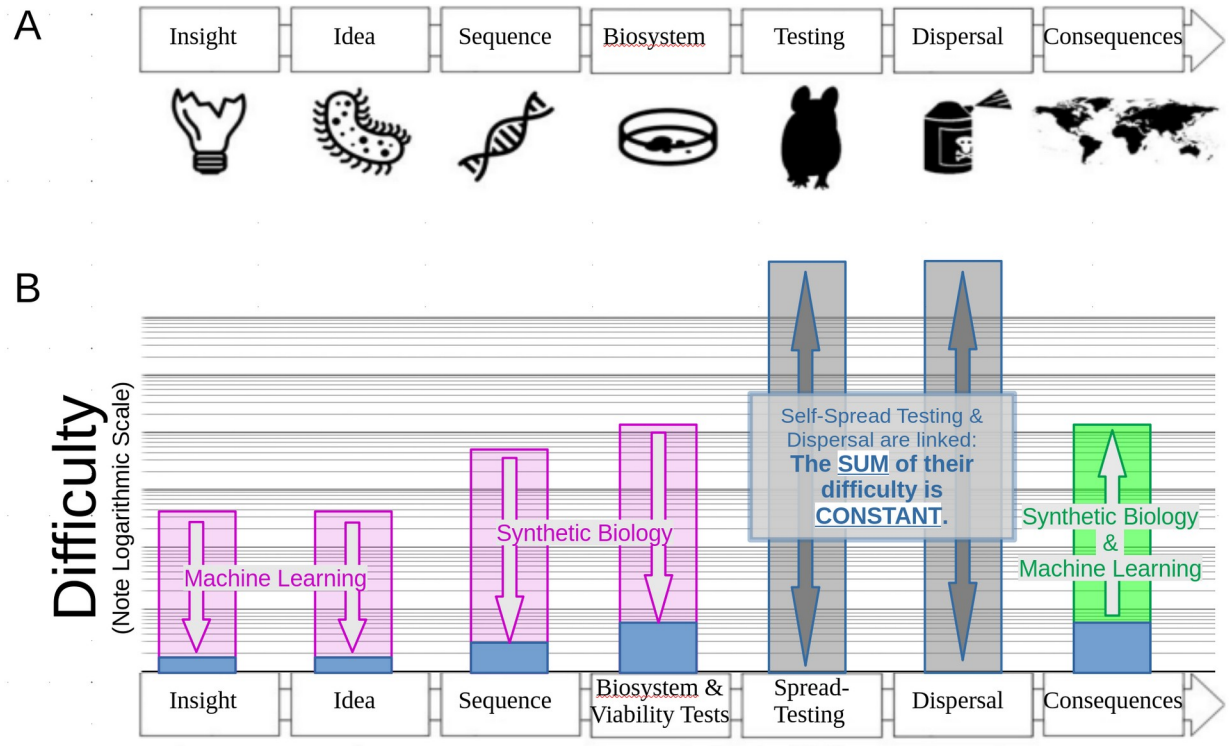


Figure 1 Caption:
A. In 2020, Sandberg and Nelson wrote about a “biorisk chain” that recapitulated and specialized the concept of technology readiness levels to the application of dangerous biotechnology such as the deliberate weaponization of an engineered microorganism. Their figure is reprinted here with permission.
B. That biorisk chain, modified to explicitly recognize spread-testing, as opposed to viability testing. as a central step, is paired with a proposed characterization of the difficulty of bringing such a hypothetical agent further down the risk chain. These difficulties are presented both before and after the introduction of recent technologies that have have revolutionized biotechnological capabilities in recent years. Notably the revolution in biotechnology is not

strictly enabling to the offender in the early steps in the biorisk chain, but also enables the defender later in the risk chain. However, the dominant barrier to an offender moving down the biorisk chain, that is to say engaging in a research and development effort to produce a biothreat agent, is the paired problems of spread-testing and dispersal. The sum of the difficult of these steps must be constant, as for every unit of spread that the agent is incapable of doing itself, further dispersal by the actor must must make up the difference for the same result and vice-versa. Further, because the limiting factor on spread testing is data collection of from the real world environment that the agent is meant to spread in, and the limiting factor on dispersal is logistic factors as in chemical agents, neither are affected by revolutionary technologies in biotechnology not artificial intelligence.

The reason spread-testing has not previously been perceived as a the defining stage of difficulty in the biorisk chain (see Figure 1), eclipsing all others, is that, until recently, the difficulties associated with the preceding steps in the risk chain were so high as to deter contemplation of the practical difficulties beyond them. With the advance of synthetic biology enabled by bioinformatic inferences on ‘omics data, the perception of these prior barriers at earlier stages of the risk chain has receded.

Spread testing represents a difficulty tens of thousands or even tens of millions of times harder than prior steps in the biorisk chain. This is a direct consequence of the fact that the spread of an agent is idiosyncratic to the particulars of the individual combination of the agent's traits and to the parts of the environment it interacts with. This eliminates any potential for spread to have anything like a one-size fits-all solution that can be solved once and then applied universally. Rather, a self-spreading agent is necessarily a bespoke product with biological and genetic traits that are not modular. Further, attempting to avoid needing to do spread testing is ‘robbing Peter to pay Paul’ as for every unit of effort evaded in designing the spread dynamics of the agent, the actor must expend a reciprocal unit of effort in ensuring that the agent reaches its intended targets via some method other than natural biological spread.

It is of course possible that spread testing can be completely eliminated by placing the burden for this functionality entirely upon delivery mechanism, but if that is the decision of the actor, then they have functionally just created a chemical weapon. Presumably, for any given actor's goals there is probably some middle ground in expenditure of effort in spread testing and delivery that represents minimum effort, but again that will be idiosyncratic to both the environment, the means of delivery available to the agent and the goals that deploying it are meant to serve once again eliminating any one-size fits-all spread solution.

Similarly, some will argue that certain classes of actor will attempt to avoid the inherent difficulty of spread testing by simply releasing agents that have not had any sort of spread-dynamic validation. The threat of this sort of release is, however, very low so long as the agent is also truly novel. Those who would make this objection imagine that such agents would simply spread slower as a consequence of not being optimized for spread. This is an expectation based upon the very limited experience that humans have at designing whole organisms with synthetic biology. However, that experience is not as informative as it seems since, first most genetically altered organisms can in fact not survive and thrive in the wild[23], [24], and second all such organisms have been either chimeras of non-novel organisms that already have the capacity to spread in their ecological niches, or near copies of such non-novel organisms with very subtle alterations. The difficulties of spread-testing by release is further explored below in the section concerning technological development cycles.

Rather, it is likely that any biological agent developed without extensive spread testing, is either mostly or entirely not novel, or fails to spread in the environment *at all*. A threat-aware understanding

of biological novelty explores how even subtle changes to an organism of seemingly validated spread characteristics renders the validity of those characteristics questionable.

Threat-Aware Understanding of Novelty and Biotechnological Threats

For most non-novel biological agents, a biotechnological acquisition process is vastly more difficult than acquiring the agent from the wild as these agents naturally emerge and re-emerge around the world.[25], [26] Thus, while non-novel biological threat agents can be either technological in origin or naturally emergent, nearly all bio*technological* threat agents can be anticipated to be novel in design and properties. Or reasoned from the other direction, if they are functionally equivalent to a non-novel agent the actor would be able to acquire that non-novel agent much more simply than engineer a novel one. Thus, is all biological agents can be classified as novel or not, and those agents of technological origin are predominantly only a subset of the novel classification, but there are also novel agents that arise naturally.

The principle exception to the correlation that biotechnological agents are always novel agents is smallpox, which is not novel, and yet is only plausibly available through the biotechnological synthesis.[27] As such, smallpox provides a useful intellectual probe to consider the contribution of novelty to the magnitude of danger represented by a biological agent. If smallpox were to re-emerge, the threat would be met by a wealth of human knowledge. There would be known vaccines, therapeutics, diagnostics, modes of spread, standards of personal protective equipment, and epidemiological models. Most importantly, public health officials and policy makers would know, not suspect but actually know, that it was a crisis of global import from the moment the first patients were correctly diagnosed. They therefore certainly could, and probably actually would, choose to take decisive action while the disease might still be containable. This vast knowledge of smallpox's biological properties is what makes it a non-novel agent and simultaneously makes a reemergence likely solvable.

Contrast that scenario with the sequence of events when COVID emerged. Minus border closings, for months very little was done, for months more diagnostics tests were unavailable or in very short supply and of uncertain reliability. It was many months into the pandemic before it was clear if public masking was of any value. Longer still before it was broadly accepted by medical and public health officials that symptom-less infection and spread was a significant epidemiological phenomenon. It was well past a year into the pandemic before it was broadly accepted that the protection of prior exposure faded with time and thus that reinfection was possible and even likely. Not knowing these things was a major contributing factor to what made COVID both novel and such a large danger.

It should not be inferred that a non-novel agent is not dangerous at all. However, it is certainly *significantly less dangerous* than it would have been if we were ignorant of its properties. Novelty of a biological agent understood in the above terms is a matter of degrees. Moreover, the degree of novelty is determined by how much we do not know about the agent's phenotypic properties not by how recently it has been known to exist. There are some biological agents which have been known to exist for many years, and yet remain novel by this threat-aware understanding simply because they are understudied. Similarly, in characterizing the novelty and thus threat of a biotechnological agent, we can ask to what degree it has properties altered from known agents. An immune evasive variant of a pathogen, for instance, is relatively novel to the non-evasive variant even if some knowledge of that non-evasive variant is still valid. This is a functional or phenotypic conception of novelty, not a genotypic one although it is presumed that any sufficiently large phenotypic alteration of a pre-existing agent will be recapitulated in its genetic sequence.

This threat-aware understanding of novelty doesn't just change our perspective of how the defenders of public health perceive the threats of biotechnology agents. It also reveals a design tension

that the attacker designer of a biotechnology threat agent experiences: The designer wants a novel agent, because novelty makes the agent more effective by the ignorance of the defender to its properties. However, if the agent is novel, then the spread properties of the agent are unknown to the attacker also, thus necessitating spread testing, because the spread dynamics of a novel agent can not be computed from first principles as discussed in above int he section on spread.

Some might argue that this tension is released by the modularity of different genes: that an attacker might add a toxicity gene or pathnogenicity island to an agent of known spread dynamics to generate a synthetic agent that was novel in pathnogenicity and yet non-novel in its spreading properties. This objection, however, ignores the true complexity and delicacy of spread as a biological phenomenon. Imagine COVID but more pathogenic; likely, such a pathogen would, while quite deadly to a few, not be pandemic-capable because the extremely low symptomatic stage of the disease which was crucial for COVID's ability to transmit would have been more severe as a consequence of the disease being more severe generally. Thus, infectious patients would be at home convalescing or in hospitals and not going about their business in public. It is worth noting that this reveals the simplicity of some measures of transmission that attempt to reduce the spread of a disease to simple numbers such as particle count and size. A disease might be more transmissible by such measures, but that does not mean that it is in fact more transmitted under real world conditions.

The ability to spread in a germ-theory-of-disease-aware civilization is a subtle balance between symptoms severe enough to achieve any transmission and weak enough so as to not alter behaviors leading to inhibited transmission. Finding that that balance is in direct tension from the a design constraint from the novel agent's designer to achieve high impact from the agent's deployment. Because of this tension, one simply CAN NOT engineer spread as a stand alone module independent of the other biological properties of the agent.

The Technology Development Cycle and Simultaneous Engineering of Spread and Novelty of a Biotechnological Agents

Technology development is a well understood phenomenon, and biotechnology is not an exception to that. Crucially, as they are developed, all technologies go through what is referred to as the Design --> Build --> Test --> Learn cycle[28], [29] through many iterations as they slowly progress up the technology readiness scale[30], [31] from concepts to fully realized products. (The biorisk chain of Sandberg and Nelson recapitulates and specializes the thinking of the technology readiness levels to the specific case of biotechnological threat agents, see Figure 1A).

An awareness of these principles of technology development alters our understanding of what a biotechnological threat R&D process would have to be capable of when also considered along side both the importance and difficulty of spread testing of biotechnological agents, and also the design tensions implicit on such a biotechnological agent given a threat aware understanding of novelty.

The first consequence of this confluence of technology development, novelty, and spread is that it alters *who* the actor of a biotechnological threat is properly considered to be. This work has focused upon the 'actor' as the biotechnological threat agent's *developer* which need not be the same party that deploys the agent. While the actor is more typically considered the deploying party, for a novel agent that can not be sourced except by through a technological resource and development pipeline, that deploying party is only as potent as the developing party that supplies them with the agent.

Second, accidental development of a high consequence biotechnological agent is vanishingly unlikely. A technology development effort must cycle through Design, to Build, to Test, to Learn many times to achieve even a working prototype much less a finalized product. One can go from Design to Build, to Test, to Learn by accident just once. But the transition from Learn back to Design, even just once to say nothing of many times over many iterations of the cycle, *requires* intention by definition.

Third, the difficulty of spread testing becomes magnified by the iterative nature of technology development. It is not enough to validate the spread dynamics of a biological agent once, rather it must be done over and over again for every iteration of the Design, Build, Test, Learn cycle. This is all the more true when one considers that the location of Spread Testing on the modified biorisk chain (see Figure 1B) is late in the biorisk chain. The delicate balance that the biological spread phenomenon is perched upon means any slight modification of any design element might alter that balance of the agent and thus requires yet another round of expensive spread testing.

Fourth, an important detail of how development cycles work is that they are necessarily *iterative;* consequently, they must be performed in series not in parallel (that is each cycle of Design --> Build --> Test --> Learn must wait for and be informed by the results of the preceding cycle). This has profound implications for a malicious actor trying to save resources by spread-testing through release of their prototype agents into the actual environment, rather than trying to recapitulate enough of the many complexities of that environment into a proxy spread-testing experimental setup. In addition to the technical problems causing such a strategy to likely fail outright, as discussed above in the section on spread, the actor would be forced to contend with a dilemma of paired downsides of inherent in environmental release: (1) They risk detection by the authorities who can also monitor the environment. (2) They must observe and qualify the success or failure of each iterative version of the developing biotechnological agent. These two downsides are in tension with one another: If the agent's effects are highly visible and thus its spread is easy to detect and identify in the environment, then it is easy to detect and identify by the authorities too. Conversely, if the agent's effects are challenging to detect and identify, then the actor must expend significant resources to do so, and the actor has no longer succeeded in evading the costs of spread testing. Further, this tension is true by degrees, consequently for every unit of evading one of these downsides, they invoke the other to the same degree making the sum of the downsides constant.

Threat Model:

These explorations of the natures of spread, novelty, and technology development allow for the synthesis a new threat model by asking the question: 'What kind of *actor* could meet the rigors of spread testing, a novel agent, over the many development cycles needed to bring an agent down the entire biorisk chain?'

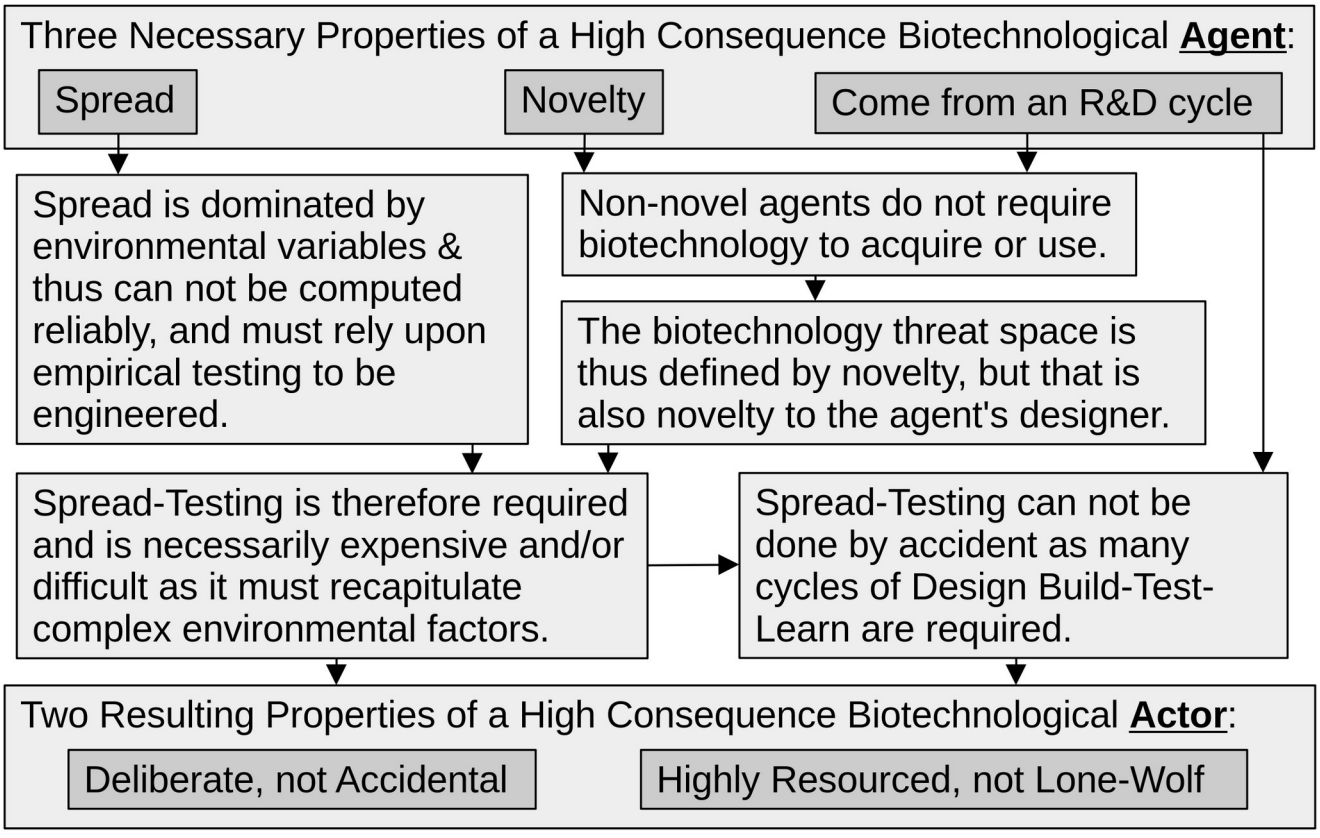
The biotechnological adversary must be (1) Intentionally motivated to do harm.[4] (2). Highly resourced or they would not be able to afford multiple rounds of spread testing through a biotechnological R&D pipeline. One can imagine them choosing to evade both the expense of spread testing and the design tension implicit in novelty by using a known agent of already known spread dynamics, but in that case, while they remain a biological actor, they are no longer a bio*technological* actor. Further, the consequences of their action is intrinsically mitigated by, and in proportion to, the same lack of novelty that makes them so.

4 Because of recent speculation of a lab-origin of COVID, it is worth noting that this understanding of biotechnology threats is bounded by a concept of biotechnology that is intentional-design-centered. An accidental lab leak of a pathogen that is naturally occurring and was never actively 'designed' is not properly a "biotechnology threat", but rather a laboratory safety threat. The concept of passage as a design approach might seem, upon first glance, to be a gray-zone allowing for accidental "design". This however, is not the case. If the passage system recapitulates the complexities of the environment than an agent is might spread in, that is, it is not just passage in a strictly laboratory setting such as tissue culture, but spread passage in an experimental set up meant to model the wild environment, then it starts to have the same properties and difficulties in terms of expense, intent, and intelligence footprint as a true spread testing endeavor, and concordantly such a spread-passage endeavor would have to be set up intentionally and at great cost.

Notably, very small organizations, disgruntled individuals, and lone-wolf actors do NOT have the resources to meet the second criteria of this threat model. Nation-state actors on the other hand, broadly avoid biological weapons as there are easier ways to make war. Potentially, there is a sweat-spot of organizational size and motivation to do harm that is big enough to resource a spread testing effort, small enough to fly under the radar, and malevolent enough to desire such novel spreading biotechnological agents. Insofar as non-technological bio-terrorism recapitulates the dynamics of the motivations and methods of biotechnological terrorist this suggests that ideologically motivated small organizations capable of focusing the resources of many members such as Aum Shinrikyo[32] or Rajneeshpuram[33] remain the most relevant potential actors for biotechnological threats.

The advance of AI technologies and the rapid growth of biotechnological capabilities demands that any grand unified biotechnology threat model must consider how further advances will alter the landscape. This model achieves a degree of future-proofness because it is based upon the underlying fact that spread dynamics are first and foremost dictated by environmental factors, and that the environment is intrinsically difficult to model, not because of a lack of understanding or a lack of compute power, but because of a lack of sufficient data collection. Further, chaos theory and the refractory nature of many systems past a certain time threshold into the future suggests that sufficient data collection to predictively model such complex dynamic wild ecosystems, and movement of epidemics and invasive biological agents or genes inside them past a similar time threshold likely has hard operational limits no matter how much data is collected. Figure 2 contains a summary of the logic of the threat model.

Figure 2



Discussion

Biotechnology is Revealed as Profoundly Defense Dominant

There is a concept commonly used in discussing threat models of technologies called the "offense-defense balance".[34] This concept recognizes the existence of the "dual use dilemma"[35], which is to say that any given technology can be used for both constructive and destructive purposes. Biotechnology, as discussed by Sandberg and Nelson[1], is perceived as relatively offense-dominant. This conclusion was arrived at as a consequence of the recognition that the early stages of the biorisk chain no longer required huge teams, high skill levels, or massive budgets. That observation coupled with the observation that biological agents could have very high consequences lead to the concern that the number of potential actors capable of designing building and releasing high consequence biological agents was very high. These hypothesized actors could include lone wolf actors with a-typical motivations and minimal resources, such as disgruntled graduate students, or garage tinkerers with more talent than sense. They might, individually, represent a low likelihood of danger, but collectively a high danger as a group because there could be so many more of them then traditional actors.

However, the threat model proposed above suggests that the low likelihood of accidental or low resources actors achieving a high consequence novel and spreading agent due to the massively difficult and necessary step of spread testing in arriving at such an agent, dramatically lowers the number of actors of concern. This reduction in the number of actors capable of the necessary R&D pipeline dramatically lower the anticipated total magnitude of biotechnology based threats now and in the future.

Not only does the reduced number of actors change the offense-defense balance towards defense, the threat model reveals that it is easier for defenders to use biotechnology to defend against threats than it is for attackers to use the same biotechnology offensively. There are two principle reasons for this:

1. Most biotechnological defense or robustness measures need not spread on their own. Indeed, because of ethical restrictions requiring informed consent, self-spreading countermeasures, such as infectious vaccines, would be intrinsically unappealing to most defenders. Without self-spread as a necessary or even appealing feature of most defense oriented biotechnology, the difficult and expensive step of spread testing, required on the offensive side of the threat-model, is bypassed for most defensive purposes.
2. Legitimate actors engaged in constructive endeavors, unlike most attackers, have the luxury of planned ongoing oversight and tuning of their interventions. While relevant to all defensive interventions, this is especially important to those interventions that DO self-spread such as proposed a malaria-control gene drive.[36], [37] A legitimate actor might choose to release such a gene drive for legitimate reasons. But, where a terrorist or attacking party might do something similar and then be forced to retreat and hope that the complex emergent behaviors of the invading biotech agent in the wild environment unfold as planned, the legitimate actor can overtly monitor and modify the intervention in an ongoing basis indefinitely. That is, the legitimate actor need not intervene in a fire-and-forget manner. This makes a huge difference in the feasibility of such an intervention because down-stream effects, for the legitimate actor only, need not be perfectly predicted at deploy-time. Rather the legitimate actor merely needs to be able to refine the intervention at a rate faster than unanticipated dynamics emerge.

The offense-defense balance is anticipated to only shift further in favor of the defensive uses of biotechnology with time. This is a function of the biorisk chain difficulties explored in Figure 1B. Note that biotechnology reduces barriers to early steps in the biorisk chain but also increases the difficulty of achieving high consequences because of the defenders utilize that same technologies. That

capacity of defenders to utilize more powerful defensive technological tools has not obvious upper limit. But the capacity of biotechnology to remove barriers for the attacker early in the risk chain has a necessary limit; the difficulty of those steps in the risk chain can't be reduced below zero. Thus in the long run, the biosecurity solution to biotechnology is more biotechnology. Indeed, biosecurity policies that slow the adoption and advance of biotechnology artificially preserve and prolong a period of relative vulnerability in which defensive uses of biotechnology have yet to fully dominate the security equation. For this reason, it is important to reconsider biosecurity policy in light of this threat model.

Implications for Biosecurity Policy and Future Study

The consequences for policy from this threat model and the three components that make it up, (that spread testing is limiting, that novelty informs threat, and that R&D rules apply) are numerous and significant.

First, agricultural environments are, like laboratories, designed to exacting reproducible specifications and specifically designed to manage or exclude confounding outside factors (weeds, pests, predators, lack of nutrients, variations in weather, etc). Such regimented farming environments therefore do not represent the intractable modeling situation that wild ecosystems do, and spread testing for them would be expected to be much easier. Farms also represent relatively fragile mono-cultures much of the time. If a class of biothreat agent escapes the ‘grand unified threat model of biotechnology’ presented here, it is likely to be inside the realm of agricultural agents. More work identifying areas of special vulnerability, such as agriculture, is needed, as is work to characterize the threats that could be most damaging in such vulnerable spaces so that countermeasures and robustness measures can be put in place.

Second, access control on the biotechnology tools that allow for the design and building of biotechnological systems is unlikely to yield significant benefits to security. The Design and Build segments of the R&D cycle are not the limiting factor of biotechnological threat agent development in terms of total difficulty regardless of whether access control measures are in place or not. That is, any organization capable of dedicating sufficient resources to spread testing will have the relatively trivial resources necessary to wholly bypass access control measures such as, for example, synthetic DNA screening. These access control measures only ever made sense when considered as a counter to low resourced or non-deliberate actors which the threat model reveals were never a meaningful threat to begin with. De-emphasizing access control has the additional advantage that it reduces barriers to legitimate research. This is especially important when one remembers that it is the legitimate research community that will develop any countermeasures to a biological threat, so reducing barriers to legitimate research directly and cumulatively aids in biosecurity.

Third, spread testing, as has been discussed above, is necessarily empirical and not computational, and involves either carefully constructed proxies, such as an animal model, or actual testing in the environment the agent is meant to be deployed in. If the method of proxies is used, it is an expensive large endeavor which by virtue of that expense has an intelligence footprint. Conversely, if the R&D of the agent uses spread-testing in the environment, it has by virtue of the environmental exposure, its own environmental footprint which might be detected by suitable ecological monitoring which should also be considered a form of intelligence gathering. Either way, these footprints combined with the unavoidable and rate-limiting nature of spread-testing in any high impact biotechnological agent's development pipeline, makes policy to enable law enforcement and intelligence agencies to detect a bioagent R&D efforts at the spread-testing stage the most efficacious biosecurity strategy. Maximizing intelligence gathering and analysis is in keeping with de-emphasizing access control measures; one wants to incentivize malign actors to use legitimate suppliers, rather than going off the grid to source their needs, as that increases the detectable footprint of their development

pipelines. Regardless, more research is needed to characterize the exact nature of the economic and environmental footprints of spread testing, what intelligence gathering methods would be most efficacious at detecting those spread testing regimes, and what levels of resources would be needed by the actors engaged in them.

Acknowledgments

The author would like to acknowledge and thank Gigi Gronvall, Lane Warmbrod, Anders Sandberg, Cassidy Nelson, Jonas Sandbrink, Richard Bruns, Matt Watson, Rebecca Montague, and Thomas Houfek for patiently listening to and critiquing numerous versions of this theory. The author would also like to thank Kevin Esvelt for numerous stimulating and cordial disagreements on related matters. This work was supported by grants from Open Philanthropy.

References

[1] A. Sandberg and C. Nelson, “Who Should We Fear More: Biohackers, Disgruntled Postdocs, or Bad Governments? A Simple Risk Chain Model of Biorisk,” *Health Secur.*, vol. 18, no. 3, pp. 155–163, Jun. 2020, doi: 10.1089/hs.2019.0115.

[2] N. D. Connell, “The challenge of global catastrophic biological risks,” *Health Secur.*, vol. 15, no. 4, pp. 345–346, 2017.

[3] N. Bostrom, “Existential risks: Analyzing human extinction scenarios and related hazards,” *J. Evol. Technol.*, vol. 9, 2002.

[4] L. G. W. Christopher, L. T. J. Cieslak, J. A. Pavlin, and E. M. Eitzen, “Biological warfare: a historical perspective,” *Jama*, vol. 278, no. 5, pp. 412–417, 1997.

[5] W. Tan *et al.*, “A Novel Coronavirus Genome Identified in a Cluster of Pneumonia Cases — Wuhan, China 2019−2020,” *China CDC Wkly.*, vol. 2, no. 4, pp. 61–62, 2020, doi: 10.46234/ccdcw2020.017.

[6] J. C. Venter, H. O. Smith, and M. D. Adams, “The Sequence of the Human Genome,” *Clin. Chem.*, vol. 61, no. 9, pp. 1207–1208, Sep. 2015, doi: 10.1373/clinchem.2014.237016.

[7] M. Apuzzo, S. Gebrekidan, and D. D. Kirkpatrick, “How the world missed COVID-19’s silent spread,” *N. Y. Times*, vol. 27, 2020.

[8] J. Giesecke, “The invisible pandemic,” *The Lancet*, vol. 395, no. 10238, p. e98, 2020.

[9] A. L. Rasmussen and S. V. Popescu, “SARS-CoV-2 transmission without symptoms,” *Science*, vol. 371, no. 6535, pp. 1206–1207, 2021.

[10] H. Deng, X. Yan, and L. Yuan, “Human genetic basis of coronavirus disease 2019,” *Signal Transduct. Target. Ther.*, vol. 6, no. 1, p. 344, 2021.

[11] M. E. Niemi, M. J. Daly, and A. Ganna, “The human genetic epidemiology of COVID-19,” *Nat. Rev. Genet.*, vol. 23, no. 9, pp. 533–546, 2022.

[12] F. B. Agusto *et al.*, “To isolate or not to isolate: The impact of changing behavior on COVID-19 transmission,” *BMC Public Health*, vol. 22, no. 1, pp. 1–20, 2022.

[13] H. Brüssow and S. Zuber, “Can a combination of vaccination and face mask wearing contain the COVID-19 pandemic?,” *Microb. Biotechnol.*, vol. 15, no. 3, pp. 721–737, 2022.

[14] T. Denayer, T. Stöhr, and M. V. Roy, “Animal models in translational medicine: Validation and prediction,” *Eur. J. Mol. Clin. Med.*, vol. 2, no. 1, p. 5, Aug. 2014, doi: 10.1016/j.nhtm.2014.08.001.

[15] C. Muñoz-Fontela *et al.*, “Animal models for COVID-19,” *Nature*, vol. 586, no. 7830, pp. 509–515, Oct. 2020, doi: 10.1038/s41586-020-2787-6.

[16] R. A. Goldstein, “Reality and models: difficulties associated with applying general ecological models to specific situations,” in *Mathematical models in biological discovery*, Springer, 1977, pp. 207–216.

[17] J. J. Bull, C. H. Remien, and S. M. Krone, “Gene-drive-mediated extinction is thwarted by population structure and evolution of sib mating,” *Evol. Med. Public Health*, vol. 2019, no. 1, pp. 66–81, 2019.

[18] J. M. Marshall and O. S. Akbari, “Can CRISPR-Based Gene Drive Be Confined in the Wild? A Question for Molecular and Population Biology,” *ACS Chem. Biol.*, vol. 13, no. 2, pp. 424–430, Feb. 2018, doi: 10.1021/acschembio.7b00923.

[19] G. A. Backus and J. A. Delborne, “Threshold-dependent gene drives in the wild: spread, controllability, and ecological uncertainty,” *BioScience*, vol. 69, no. 11, pp. 900–907, 2019.

[20] Y. M. Buckley, “Invasion ecology: Unpredictable arms race in a jam jar,” *Nat. Ecol. Evol.*, vol. 1, no. 1, p. 0028, 2017.

[21] R. DeFries and H. Nagendra, “Ecosystem management as a wicked problem,” *Science*, vol. 356, no. 6335, pp. 265–270, 2017.
[22] J. J. Kay and E. Schneider, “Embracing complexity the challenge of the ecosystem approach,” in *Perspectives on ecological integrity*, Springer, 1995, pp. 49–59.
[23] L. Lambrechts, J. C. Koella, and C. Boete, “Can transgenic mosquitoes afford the fitness cost?,” *Trends Parasitol.*, vol. 24, no. 1, pp. 4–7, 2008.
[24] E. Jenczewski, J. Ronfort, and A.-M. Chèvre, “Crop-to-wild gene flow, introgression and possible fitness effects of transgenes,” *Environ. Biosafety Res.*, vol. 2, no. 1, pp. 9–24, 2003.
[25] S. A. Walper *et al.*, “Detecting biothreat agents: From current diagnostics to developing sensor technologies,” *ACS Sens.*, vol. 3, no. 10, pp. 1894–2024, 2018.
[26] H. D. Marston, G. K. Folkers, D. M. Morens, and A. S. Fauci, “Emerging viral diseases: confronting threats with new technologies,” *Sci. Transl. Med.*, vol. 6, no. 253, pp. 253ps10-253ps10, 2014.
[27] G. D. Koblentz, “The de novo synthesis of horsepox virus: implications for biosecurity and recommendations for preventing the reemergence of smallpox,” *Health Secur.*, vol. 15, no. 6, pp. 620–628, 2017.
[28] P. Opgenorth *et al.*, “Lessons from two design–build–test–learn cycles of dodecanol production in Escherichia coli aided by machine learning,” *ACS Synth. Biol.*, vol. 8, no. 6, pp. 1337–1351, 2019.
[29] P. Carbonell *et al.*, “An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals,” *Commun. Biol.*, vol. 1, no. 1, p. 66, 2018.
[30] J. C. Mankins, “Technology readiness levels,” *White Pap. April*, vol. 6, no. 1995, p. 1995, 1995.
[31] P. V. Almeida, L. M. Gando-Ferreira, and M. J. Quina, “Biorefinery perspective for industrial potato peel management: technology readiness level and economic assessment,” *J. Environ. Chem. Eng.*, p. 110049, 2023.
[32] A. T. Tu, “Aum Shinrikyo’s chemical and biological weapons: more than sarin,” *Forensic Sci Rev*, vol. 26, no. 2, pp. 115–20, 2014.
[33] J. Thomas *et al.*, “A large community outbreak of salmonellosis caused by intentional contamination of restaurant salad bars,” *Jama*, vol. 278, no. 5, pp. 389–395, 1997.
[34] C. L. Glaser and C. Kaufmann, “What is the offense-defense balance and can we measure it?,” *Int. Secur.*, vol. 22, no. 4, pp. 44–82, 1998.
[35] J. Rath, M. Ischi, and D. Perkins, “Evolution of different dual-use concepts in international and national law and its implications on research ethics and governance,” *Sci. Eng. Ethics*, vol. 20, pp. 769–790, 2014.
[36] J. L. Frieß, C. R. Lalyer, B. Giese, S. Simon, and M. Otto, “Review of gene drive modelling and implications for risk assessment of gene drive organisms,” *Ecol. Model.*, vol. 478, p. 110285, 2023.
[37] K. L. Warmbrod, A. Kobokovich, R. West, G. Ray, M. Trotochaud, and M. Montague, “Gene Drives: Pursuing Opportunities, Minimizing Risk.” The Center for Health Security, Johns Hopkins, May 20, 2020. [Online]. Available: https://centerforhealthsecurity.org/2020/new-report-from-johns-hopkins-center-for-health-security-gene-drives-pursuing-opportunities-minimizing-risk-0