The Accelerating Pace of the Democratization of Biotechnology

Shawn S. Jackson[1*], Louise E. Sumner[1], Christian H. Garnier[1], Casey Basham[1], Landy T. Sun[1], Peter L. Simone[1], Danielle S. Gardner[1], and Rocco J. Casagrande[1]

[1] Gryphon Scientific, LLC, Takoma Park, MD, USA

* Corresponding Author: Shawn S. Jackson

Gryphon Scientific, LLC

6930 Carroll Avenue, Suite 900

Takoma Park, MD 20912

Tel: +1+240-485-2542

froggi@gryphonscientific.com

Email addresses of authors:

SSJ: froggi@gryphonscientific.com

LES: louise@gryphonscientific.com

CHG: cgarnier@gryphonscientific.com

CB: cbasham@gryphonscientific.com

LTS: lsun@gryphonscientific.com

PS: plssimone@gmail.com

DSG: dgardner@gryphonscientific.com

RJC: rocco@gryphonscientific.com

Understanding the speed at which new biotechnologies develop from activities requiring significant educational and financial investments into those requiring far less resources is important for developing regulatory and security policies and practices. Moreover, such understanding provides insight on how these policies and practices, necessary to ensure that cutting edge research is practiced responsibly, influence the biotechnology economy. Here, an analytic approach was developed to generate supportable estimates of the pace of advancement of biotechnologies. When extrapolated into the future, these estimates predict timeframes for the democratization of novel biotechnologies, that is, the speed at which they can be expected to transition from the hands of a few well-resourced specialists into those of individuals with relatively low levels of technical skill and financial resources. Our assessment provides evidence that novel technologies can currently make this transition in less than four and half years from their discovery, and by the close of the next decade, could transition in less than three and a half years.

Historically, new techniques and devices used in biotechnology have spread quickly throughout molecular and cellular biology research fields. With ensuing technological refinements, methods, reagents, and equipment become cheaper and more widely available.[1, 2] Eventually, some biotechnologies reach the point where they no longer require specialized skills or substantial financial resources to perform, and can be successfully used by individuals with relatively low levels of technical skill and financial resources. Some biotechnologies have such broad applications that demand supports the development of service industries. Through the study of a variety of sources, we followed the development of new biotechnologies with the aim of understanding the pace at which new biotechnologies spread throughout the field and progress from activities requiring significant technical training and financial investments to those requiring far less resources. Data collected were extrapolated to estimate how quickly new biotechnologies could spread in the future. From these estimates we determined that the time required for the democratization of biotechnologies has been decreasing, and moreover, that the pace of this decrease, in the time required for techniques to become widely accessible, has been accelerating, as the technique's inception date has become more recent. Although the future of advances in biotechnology cannot be predicted with certainty, our study indicates that the potential for the use of biotechnologies by non-experts is progressing at a rapid pace.

To generate widely-applicable estimates of timelines for the development and spread of biotechnologies, we investigated the histories of 22 biotechnologies selected for their widespread influence on molecular and cellular biology laboratory practices (Table 1). Chronologies for each biotechnology were defined using open-source information retrieved via internet-based searches, including published literature, government reports, business and industrial histories, and on-line university biology course materials, among others. From these sources, milestones that indicated the spread of each biotechnology were noted, beginning with the first demonstration of the technique. Then, the spread of the technique from the initiating laboratory to other laboratories, the extension of the technique's applicability to additional experimental systems, and the advent and increases in the accessibility of commercially-available reagents (especially "kits") and equipment for the technology were recorded (see Table 1 and Supplementary Table 1 for milestone examples).

Using the milestone chronologies as a framework, we identified documentable evidence for each biotechnology's "spread," and evaluated the relative cost and technical skill required to successfully implement each biotechnology at each milestone, thus reflecting its accessibility to an individual with a given level of technical skill and financial resources. In so doing, our assessment of the accessibility of each biotechnology applied both objective and subjective metrics. For subjective metrics, four categories were defined to characterize both the skill and financial resource requirements needed at each stage of a biotechnology's development: very high, high, medium, and low. For example, the "invention" of a biotechnology was often the product of work of highly-resourced individuals, while exposure to the technology in educational settings such as university or even secondary school laboratories was considered to demonstrate a biotechnology's successful application by low-resourced individuals. The establishment of fee-for-service providers was generally considered to represent a biotechnology's potential application by individuals (the buyers) with low skill but not necessarily low financial resources (as the price of the service might still be high). Thus these categories encompassed biotechnologies requiring multiple investigators with advanced experience or the support of multiple laboratories for success, as well as those requiring no training beyond secondary school biology. The easing of the technical skill and financial resource requirements that occur during the democratization of a biotechnology was thus tracked. The definition of these rating categories (discussed in the online Methods) and the placement of each biotechnology's milestones within them required the development and application of a qualitative evaluation rubric based upon the expertise of the authors. Although assignments of resource need levels therefore face possible systemic errors, similarities in the predictive trends generated from both objective and subjective metrics support the assertion that both sets of data provide valuable insight.

The source information, the chronologies, and the requirements ratings for the 22 studied biotechnologies are provided in detail in the Supplementary Material. Here, we present compiled data for the subset of development milestones noted as being highly informative for assessing the pace of the development and spread of biotechnologies; these selected development milestones and the associated data are listed in Table 1. Again, note that the presented data are quantitative. For some milestones, the data reflect the authors' qualitative resource requirement categorizations; the categorizations were used as indicators for calendar year assignments.

As presented in Table 1, data were collected from biotechnologies that were initiated over the course of 75 years (1937-2012) of biological and chemical research. Once described, reproduction of these new technologies occurred with extreme rapidity – over three-quarters of the biotechnologies were replicated by other researchers within a year of their first description, and all had been replicated within 4 years. The progression of each biotechnology through the other studied milestones was considerably more variable than the early replication events, with milestones reached anywhere from one to 48 years after a biotechnology's initiation. Notably, moving from left to right across Table 1 – from "older" to "newer" biotechnologies – the reader can readily see that these durations generally decrease. In summary, the compiled data for all 22 biotechnologies show that innovations in biotechnologies spread throughout the scientific community quickly, taking a median of just one year to spread to other laboratories, and a median of just three years to be adapted for use in alternative systems (Table 1). Further progression of the biotechnology puts it in the hands of low-skilled and low-resourced individuals within 12-13 years and in the hands of medium-skilled and -resourced individuals in about half that amount of time (medians). Widely useful biotechnologies spur the formation of fee-for-service industries within 16 years from initial publication (median). As documentation of activities was not always easy to track down, errors in milestone assignments due to a lack of available evidence can be assumed to have occurred despite the authors' best efforts. Since these difficulties in obtaining evidence would invariably result in the assignation of a later, rather than earlier, milestone year, these errors in milestone assignments would have tended to bias toward longer development intervals (and a slower developmental pace).

With the chronologies and the requirement ratings in hand, we used regression analyses to formulate predictions for current and future trends in the development and spread of novel biotechnologies. Given the shapes of the plotted data (Figure 1) and the assumptions of both linear and nonlinear regression analysis, we evaluated the relationships between time and years to milestones using exponential regression analysis rather than linear, generalized linear or some other functional nonlinear regression analysis. We used available data from all 22 chronologies to define time-dependent relationships between a biotechnology's initiation year and the speed of its progression through particular development milestones. In doing so, we determined that the pace of this progression is accelerating with time. Figure 1 presents a focused view of the time-dependent relationships between a biotechnology's initiation year and the speed of its progression to low- or medium-resource requirements, and to its commercialization or entry into educational curricula . Medium-rated technical skills were those requiring a Bachelor's degree or equivalent experience for success; biotechnologies requiring no training beyond secondary school biology for success were considered to have low skill needs. Medium-rated financial requirements indicated the need to acquire some equipment, including fermenters up to 5L; biotechnologies requiring no access to an established facility, and including improvised fermenters up to 5L, were considered to have low financial needs (see Online Methods for a detailed discussion of ratings). The plots in Figure 1 present all available data from the studied biotechnologies; as not all milestones were recorded for each biotechnology, the sample size used to analyze each of the presented relationships varies as indicated in Table 2. Due to the many sources of error faced in this study, we present these data with upper and lower 95% confidence bands. In Figure 1, one can see that these confidence bands "spread wider" as the analyzed data nears the present time; this is to be expected given the many inaccuracies inherent in using historic data to predict future trends.

Using the relationships presented in Figure 1 to predict the speed at which new biotechnologies will propagate now and in the future, we found that in the near term (2020), a novel biotechnology is predicted to be accessible to a low-resourced individual in less than 4.5 years from its initial publication (Table 2). Due to the accelerating pace of biotechnology development defined by our data, this accessibility is predicted to occur more quickly in the future. For example, by 2030, newly discovered biotechnologies will be in the hands of low-resourced individuals in less than 3.5 years from the time they are first published (Table 2). The many potential sources of error faced by this study serve to explain the variable confidence in the presented estimates seen in Table 2. As noted, difficulties in obtaining evidence would have invariably resulted in the assignation of later milestone years, biasing our predictive estimates toward a slower developmental pace. The idea that these predictions may be trending slow is astonishing given the rapidity of development indicated.

Of particular interest is the predicted relationship for the formation of fee-for-service industries from novel biotechnologies. This robust relationship (R^2 = 0.81) indicates that the progression of new biotechnologies to fee-for-service companies is already very rapid, currently taking less than 2.5 years (Table 2). This relationship underpins the supposition that biotechnologies are being commercialized at an ever faster pace. Therefore, individuals who do not currently have the necessary technical or financial background to perform the techniques successfully themselves may not need to wait for further developments leading to reductions in cost or required skills. Instead, such individuals may be able to access the techniques in the near future by taking advantage of fee-for-service companies. Moreover, the line depicting our predictive relationship for initiation year and the development of fee-for-service companies for a novel biotechnology has the steepest negative slope of any relationship in Figure 1. Thus while the democratization of biotechnology as a whole is accelerating, the spawning of fee-for-service companies from novel biotechnologies is accelerating at a faster rate than reductions in either the needed skills or cost.

Finally, biotechnologies become rapidly democratized by their teaching in university laboratories; our data predicts that in general, new biotechnologies can reach such classrooms before they reach individuals at the low technical and financial resource levels (Table 2). Teaching at the university, and even the secondary school level, has become a reliable means to expose individuals to biotechnologies that they might otherwise not have access to (see Supplementary Material).

Overall, our assessment demonstrates the accelerating democratization of biotechnology, and provides clear evidence that technical knowledge can spread and equipment and reagents can be developed for widespread use within the span of just a few years. Inherent limitations of this timeline analysis are that, among novel biotechnologies, only those that progressed were included, and among recent biotechnologies, only those that progressed relatively quickly could be captured. However, since the overall trends observed hold across the entire time frame studied, this limitation is considered to be minor. The acceleration in adoption of biotechnologies with more recent invention discussed here mirrors that observed globally for technologies ranging from transportation (steam engines), electricity, and telecommunication to modern information technologies (internet) (https://www.weforum.org/agenda/2018/02/the-rising-speed-of-technological-adoption,[3]). In fact, Comin and Hobijn determined that with every decade that passes, newly invented technologies find widespread adoption more than 4 years more quickly;[3] our results focusing on biotechnologies are similar, with predictions that biotechnologies will be adopted 1 year more quickly as we move from 2020 to 2030. Continued advancement of biotechnology capabilities, accumulation of relevant research, and ease of information exchange are expected to contribute to this acceleration in democratization moving forward. Furthermore, successful new biotechnologies can quickly become ubiquitous and evolve into drivers of the biotechnology economy. One need only look at the speed with which the use of CRISPR editing has become widespread to know that this outcome is possible.[4-7] Clearly, the pace of advancement in biotechnology is progressing to a point where developments will occur with startling quickness. These results underpin the necessity of the constant review of the security implications of the democratization of powerful biotechnologies, and the proactive development of policies, oversight and guidance systems to ensure that they are leveraged responsibly by those outside the established scientific community.

Analyses such as these should be revisited often to ensure we maintain a clear grasp on the potential consequences of the capabilities afforded us by biotechnologies.

Methods

The chronological development of 22 biotechnologies was investigated (Table 1); the biotechnologies were selected for their current widespread use in fields of molecular and cellular biology. Chronologies describing milestones along the development of the techniques and their spread throughout similar and related fields of study were generated for each technique (see Supplementary Table 1 for milestone examples). Chronologies were defined using open source information retrieved via internet-based searches, including published literature, government reports, business and industrial histories, and educational institution records, among others. The details of these chronologies are provided in the Supplementary Material. An inherent limitation applies to this analysis: among recent biotechnologies, only those that progressed relatively quickly could be captured.

The accessibility of each biotechnology to a hypothetical individual were assessed by both subjective and objective metrics. Objective metrics included the durations of time taken for a new biotechnology to be marketed/commercialized (reagent/equipment available for purchase), offered as a fee-for-service by industry, and taught in university laboratory courses, Subjectively, the relative technical skill and financial resources that would be required to successfully perform each biotechnology over its development history were evaluated at each developmental milestone. The milestone chronologies were developed with the goal of reflecting the speed, cost, and technical skill required to successfully implement the technology. An individual's level of technical skill and financial resources were assumed to have a substantial influence on that individual's ability to successfully implement a biotechnology. Four categories were defined to characterize both the skill and financial resource requirements needed at each stage of a biotechnology's development: very high, high, medium, and low.

Biotechnologies requiring very high technical skills were defined as those whose success likely depended upon multiple persons with advanced experience (post-graduate work and/or industrial experience). Highly-rated technical skills required advanced experience (post-graduate work and/or industrial experience) for success and medium-rated technical skills required a Bachelor's degree or equivalent experience for success. Biotechnologies requiring no training beyond secondary school biology for success were considered to have low skill needs.

Biotechnologies requiring very high financial resources were defined as those whose success likely depended upon access to multiple existing laboratories/ facilities and the entire range of bacterial production methods. Highly-rated financial resources were required for those biotechnologies where access to an existing facility and the entire range of bacterial production methods were needed for success, while the success of biotechnologies with medium-rated financial requirements did not require access to a facility, but did require enough financial resources to acquire some equipment, including fermenters up to 5L. Biotechnologies requiring no access to or establishment of a facility, and including resources applicable to toxin production and improvised fermenters up to 5L, were considered to have low financial needs for success.

For each biotechnology, evidence for its commercialization, availability as a fee-for-service, or teaching in university laboratory courses, as well as the assignments made for technical and financial resource needs throughout the course of the development chronologies, are provided in the Supplementary Material (Supplementary Tables 2-23).

For each developmental milestone, data were plotted in terms of achievement date or resource need with respect to the initiation year of each biotechnology that met that milestone, and the resulting described relationships determined by calculating the best-fit line for the data using an exponential regression. We selected exponential regression rather than linear, generalized linear, or some other nonlinear regression analysis to examine the relationship between initiation year and the years to reach each milestone. Assessing the assumptions of linear or nonlinear regression was difficult due to the limited data available to build the models; while 22 biotechnologies were investigated, not all milestones were met for all biotechnologies. However, the normality of the log-transformed residuals of the exponential model were more normally distributed than the residuals of the linear model and the fit of the models (using R^2 as the measure of goodness of fit) were better when modeling the relationships using the exponential form.

Using a 95% confidence level, we calculated upper and lower confidence bands. These bands capture the range of confidence intervals of the lines of best-fit from each regression analysis. We expect 95% of data samples to yield a fitted line within those bands. Regression statistics of sample size, R^2, coefficient, intercept, and critical t-value are reported for each analysis. To summarize the entire set of all 22 studied biotechnologies, median values are reported (Table 1) as data that originated as the year (date) of particular occurrences were assumed to be non-normally distributed.

Acknowledgements

The authors thank Ms. Laurel MacMillan for assistance in statistical analysis and Mr. Noah Gunther for assistance in formatting. The authors also thank Dr. Jason K. Middleton at Battelle Memorial Institute for project support. Work was supported by Department of Homeland Security prime contract FA8075-14-D-0003/DO 0010 to Battelle Memorial Institute.

Author Contributions

S.S.J., L.E.S., and R.J.C. designed the analysis. S.S.J. and L.E.S. directed the data collection efforts of C.H.G., C.B., D.S.G., P.L.S. and L.T.S., and S.S.J. analyzed the data. S.S.J. and L.E.S. wrote the manuscript with input from R.J.C.

Competing Financial Interest Statement

The authors declare no competing financial interests.

References

1. Carlson, R. The pace and proliferation of biological technologies. *Biosecurity and bioterrorism : biodefense strategy, practice, and science* **1**, 203-214 (2003).
2. Carlson, R. The changing economics of DNA synthesis. *Nature biotechnology* **27**, 1091-1094 (2009).
3. Comin, D. & Hobijn, B. An Exploration of Technology Diffusion. *American Economic Review* **100**, 2031-2059 (2010).
4. Adli, M. The CRISPR tool kit for genome editing and beyond. *Nature communications* **9**, 1911 (2018).
5. Hatada, I. & Horii, T. Genome editing: A breakthrough in life science and medicine. *Endocrine journal* **63**, 105-110 (2016).
6. Khan, S. et al. CRISPR/Cas9: the Jedi against the dark empire of diseases. *Journal of biomedical science* **25**, 29 (2018).
7. Lander, E.S. The Heroes of CRISPR. *Cell* **164**, 18-28 (2016).

Figure Legends

Figure 1. Time-Dependence of the Progression of Biotechnology Development. The development histories of 22 biotechnologies (Table 1) were investigated and milestones that indicated the spread and democratization of each biotechnology were noted. Abbreviated milestone examples are presented in Table 1; the full list of milestones is presented in Supplementary Table 1, along with notional technical and financial requirements ratings. Development histories for the 22 biotechnologies are provided in the Supplementary Material and include the technical and financial requirements ratings assigned to each at the appropriate noted milestones (Supplementary Tables 2-23). Collected information was analyzed in terms of the duration of time (in years) between what was considered the initiation or invention of the biotechnology and the attainment of each milestone and/or step "down" in requirements ratings. From these data for all 22 biotechnologies, regression analyses defined time-dependent relationships between a biotechnology's initiation year and the speed of its progression through particular development milestones. The resulting predictive relationships for the duration of time required for a new biotechnology to be marketed/commercialized (a), offered as a fee-for-service (b), taught in university laboratory courses (c), require medium-rated skill (d) or financial (e) resources, or require low-rated skill (f) or financial (g) resources, are presented. Note that the relative financial and technical resources required to successfully implement each biotechnology were evaluated separately at each milestone from among four categories: very high, high, medium, and low (defined in the Methods). Panels (d)-(g) therefore represent analyses of qualitative assessments while panels (a)-(c) reflect entirely quantitative assessments. In each panel, the solid black lines are the exponential regression-predicted values for the shown plotted points while the 95% confidence bands are in dashed gray. The downward slopes of these lines indicate negative relationships between time and the years to reach the milestones. That is, the pace of these progressions are accelerating with time.

Data availability statement

Source data for Tables 1 and 2 and Figure 1 are provided in the Supplementary Material file.

Table 1. Overview of Biotechnology Development Results.

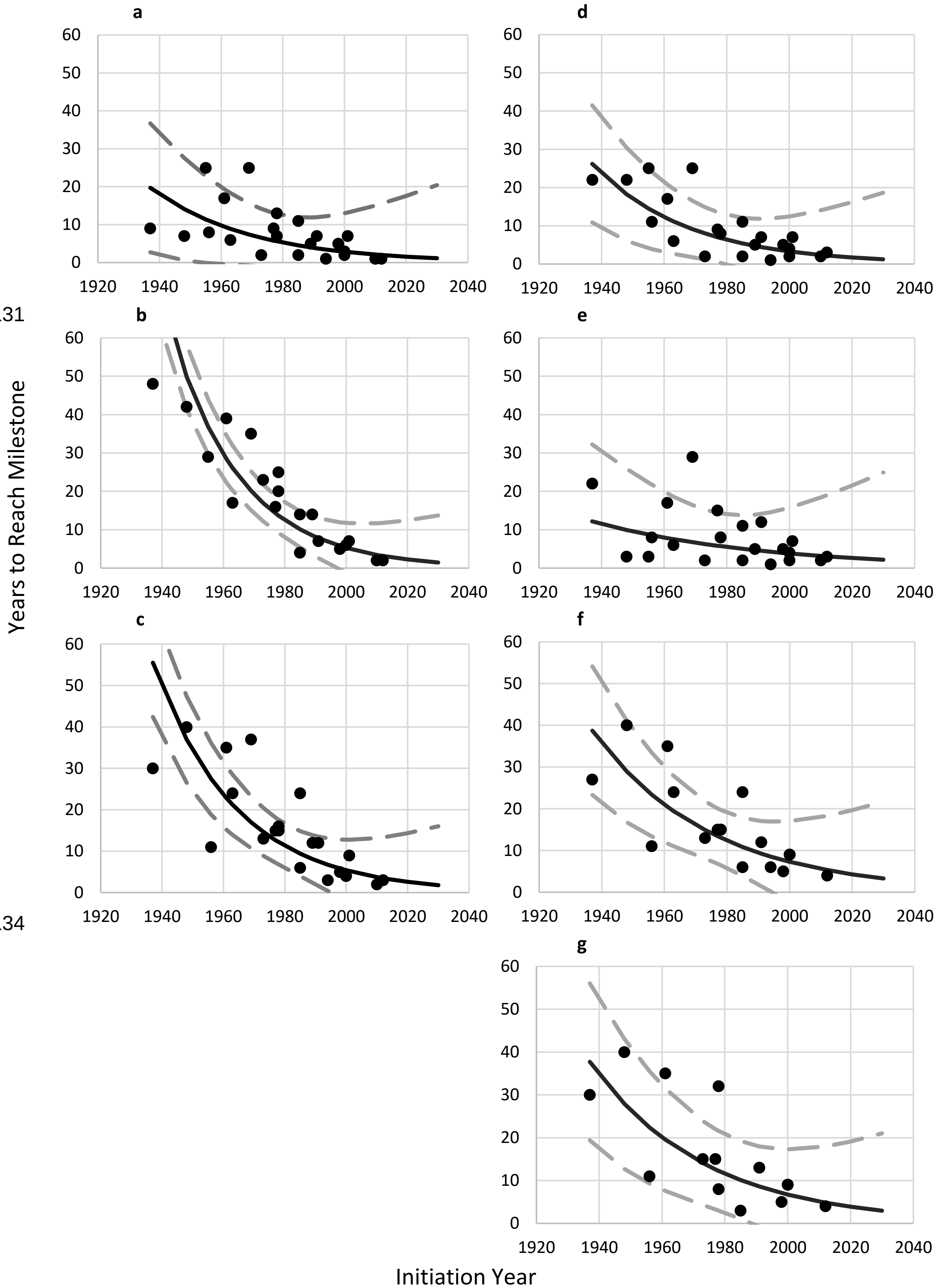
Development Milestone (Years from Initiation until Milestone)* / **Technology**	Protein purification	Mammalian cell culture	Oligonucleotide synthesis	Nucleotide (DNA and/or RNA) purification	Cell-free protein synthesis (CFPS)	Solid-phase peptide synthesis (SPSS)	*In situ* hybridization (ISH)	Molecular cloning	DNA sequencing	Site-directed mutagenesis (SDM)	Recombinant protein expression	Phage display	Polymerase chain reaction (PCR)	Yeast-2-hybrid system (Y2H)	DNA microarray	Green fluorescent protein (GFP)	RNA interference (RNAi)	Synthetic genetic circuits	Targeting induced local lesions in genomes (TILLING)	Zinc-finger nuclease (ZFN) genome editing	Transcription activator-like effector nuclease (TALEN) genome editing	Clustered regularly interspaced short palindromic repeats (CRISPR) genome editing	**Median**
“Initiation Year”	1937	1948	1955	1956	1961	1963	1969	1973	1977	1978	1978	1985	1985	1989	1991	1994	1998	2000	2000	2001	2010	2012	1981
Reproduced in other labs	1	1	1	1	0	0	0	1	1	0	2	3	1	2	4	0	0	0	0	2	1	0	1
Reproduced in other systems	4	3		1	3	7	25	1			5	5	7	5	9	1	2		3	1	1	3	3
Commercialization	9	7	25	8	17	6	25	2	9	13	7	11	2	5	7	1	5	**3**	2	7	1	1	7
Fee for service	48	42	29		39	17	35	23	16	20	25	14	4	14	7		5		6	7	2	2	16
University lab class	30	40		11	35	24	37	13	15	15	16	24	6	12	12	3	5	4	4	9	2	3	12
Low skill	27	40		11	35	24		13	15	15	15	24	6		12	6	5	9	9			4	13
Medium skill	22	22	25	11	17	6	25	2	9	8	8	11	2	5	7	1	5	2	4	7	2	3	7
Low finance	30	40		11	35			15	15	32	8		3		13		5	9	9			4	12
Medium finance	22	3		8	17	6	29	2	15	8	8	11	2	2	12	1	5	2	4	7	2	3	6

** Definitions, details, and references for milestone dates and technical and financial resource requirements are provided in the Supplementary Material. Blank cells indicate specific evidence of milestone attainment could not be found.*

Table 2. Predicted Time (Years) Required for a Biotechnology to Reach Progression Milestones in Future Years.

Progression Milestone	Statistical Support					Future Year		
	n	R^2	Coefficient	Intercept	T-value	2020	2025	2030
Medium Skill	22	0.52	-0.03	66.9	2.1	1.7	1.5	1.2
Medium Finance	22	0.18	-0.02	38.3	2.1	2.6	2.4	2.2
Low Skill	17	0.63	-0.03	54.9	2.1	4.3	3.8	3.3
Low Finance	14	0.53	-0.03	56.6	2.2	3.9	3.4	3.0
University Lab Class	21	0.68	-0.04	75.7	2.1	2.6	2.1	1.8
Commercialization	22	0.41	-0.03	62.1	2.1	1.6	1.3	1.2
Fee-for-Service	19	0.81	-0.04	87.8	2.1	2.2	1.8	1.5

Figure 1.



138