**Introduction**

Will the domestication of high technology, which we have seen marching from triumph to triumph with the advent of personal computers and GPS receivers and digital cameras, soon be extended from physical technology to biotechnology? I believe that the answer to this question is yes. Here I am bold enough to make a definite prediction. I predict that the domestication of biotechnology will dominate our lives during the next fifty years at least as much as the domestication of computers has dominated our lives during the previous fifty years.[[1]](#footnote-1)

This is Freeman Dyson writing in 2007. Dyson was one of the physicists who laid the groundwork for the transformation of physics and physics-based technology in the second half of the twentieth century. He is not alone in suggesting that biology will be to the twenty-first century what physics was to the twentieth. Over the last several decades, biology has already been transforming. The growing importance of bioinformatics – computers, data, and data-management – signals that fundamental changes have taken place in how biologists work and how they build knowledge.

But the great success and the ultimate ‘domestication’ of physics in the second half of the twentieth century came at a significant price. The discipline was overhauled to the extent that would have made it hardly recognizable to practitioners of just a few decades earlier. Physics was transformed by the need for more money, more manpower, more machines, and military involvement.[[2]](#footnote-2) Physicists struggled with their changing identity as their work became increasingly oriented towards the management of large interdisciplinary teams working at centralized institutions. What it meant to *do* physics was fundamentally transformed over the middle decades of the century – practices, institutions, reward structures, career paths, education, and the core knowledge of the discipline were transformed.

This transition did not occur without a struggle. Some physicists – and some non-physicists – worried intensely about what the advent of Big Science would mean for not only their own professional futures, but also for the future of scientific knowledge. Although it was Donald Glaser’s invention of the bubble chamber that provided the model for the first large-scale post-war physics laboratories, he was so disturbed by the new culture that he quit the discipline entirely. Glaser longed for the days of the individual scientist working diligently at his lab bench and disliked the corporate, collective world that physics had become.[[3]](#footnote-3)

Famously, in his valedictory address President Dwight Eisenhower worried about the corrupting effects that large amounts of money would have on the purity of scientifically-produced knowledge; the ‘military-industrial’ complex, he fretted, posed an immediate threat to free ideas and intellectual curiosity:

Today, the solitary inventor, tinkering in his shop, has been overshadowed by task forces of scientists in laboratories and testing fields. In the same fashion, the free university, historically the fountainhead of free ideas and scientific discovery, has experienced a revolution in the conduct of research. Partly because of the huge costs involved, a government contract becomes virtually a substitute for intellectual curiosity. For every old blackboard there are now hundreds of new electronic computers.[[4]](#footnote-4)

Alvin Weinberg, the director of Oak Ridge, the sprawling national laboratory that had helped to build the first atomic bombs, was concerned also about the effects of Big Science both on scientists and society. The effects of Big Science were felt most strongly for the individual investigators:

The research professor in a simpler day concerned himself with the substance of his science, both in research and teaching. Now, through no fault of his own, he must concern himself with many other matters. To do his research, he has to manage, even in Little Science, fairly large sums of government money. He must write justifications for his grants; he must serve on committees that recommend who should receive support, who should not; he must travel to Washington either to advise a government agency or to cajole a reluctant contract administrator. In short, the research professor must be an operator as well as a scientist.[[5]](#footnote-5)

Such changes, Weinberg argued, led to a reduced flexibility in scientific practice, compromised its “intellectual discipline,” created increasing fragmentation between the specialized branches of science, and forced difficult decisions about what kinds of science to support. Weinberg particularly criticized high-energy physics – the apotheosis of Big Science – for its failure to address problems of human welfare and technology, and its remoteness even from other branches of physics.[[6]](#footnote-6) Those who worked in less well-funded sub-disciplines, such as solid-state physics, wondered whether high-energy physics really deserved so much cash. In 1972, Philip W. Anderson argued that the “arrogance” of particle physics was causing important domains of investigation to be under-studied and under-funded; solid state physics presented an alternative to the reductionist approach of Big Science.[[7]](#footnote-7) Such arguments lay at the surface of a deeper pool of discomfort about what money, machines and the military had done to physics. Big Science forced reflection on the most appropriate ways to make physical knowledge.[[8]](#footnote-8)

Biology, as it too becomes Big Science, is currently experiencing many of the same changes that revolutionized physics mid-last century. Significantly, many of these changes are being driven by the computerization of biological work and knowledge. Eisenhower’s image of blackboards transforming into electronic computers reminds us that computers not only played a crucial role in transforming physics, but also became an important symbol of the new styles of working and knowing. Information technology – computers, networks, robotics – are playing an even more central role in transforming the practices of biology in the early twenty-first century.

What did a biology laboratory look like before computers, Big Science, and genomics? In the 1970s, the ethnographers Bruno Latour and Steve Woolgar gave a first-hand account of a neuroendocrinology lab from the 1970s. They described workers preparing assays, attending to instruments, and recording data. They described the flows of paper (mail, journal off-prints, and typed manuscripts), animals, and chemicals through the lab.[[9]](#footnote-9) No doubt there are many laboratories that still more or less conform to this description: work is organized around benches, which are filled with reagents and samples, technicians and scientists communicate and collaborate with one other primarily in face-to-face interactions around blackboards or whiteboards, and the most important things that come in and out of the laboratory are physical materials and printed papers.

But by the mid-1980s, changes were afoot. In 1991, Walter Gilbert, one of the inventors of DNA sequencing, proclaimed the coming of a ‘paradigm shift’ in molecular biology: “Molecular biology is dead – long live molecular biology!”[[10]](#footnote-10) The old ‘experimental’ paradigm was to “identify a gene by some direct experimental procedure – determined by some property of its product or otherwise related to its phenotype, to clone it, to sequence it, to make its product and to continue to work experimentally so as to seek and understanding of its function.” Gilbert worried that this “list of techniques” was making molecular biology more technology than science – it would become “pocket molecular biology” that could be performed by reading a recipe book or purchasing a kit.[[11]](#footnote-11)

His answer to this was a new paradigm in which:

all the ‘genes’ will be known (in the sense of being resident in databases available electronically), and that the starting point of a biological investigation will be theoretical. An individual scientist will begin with a theoretical conjecture, only then turning to experiment to follow or test that hypothesis.[[12]](#footnote-12)

Gilbert emphasizes the need for biologists to become computer literate, to “hook our individual computers into the worldwide network that gives us access to daily changes in the database” and thus to take advantage of the overwhelming amounts of data to “change their approach to the problem of understanding life.”[[13]](#footnote-13) Just as the Human Genome Project was gearing up, Gilbert perceived a change towards a highly data-driven and even ‘theoretical’ approach to biological problems.

But the kind of problems and challenges Gilbert identified were just the tip of the iceberg. Just like in physics, the transition to computerized, big biology required new institutions, new forms of organization, new ways of distributing resources, new forms of practice, and new kinds of knowledge.

When I ventured into a laboratory in the Department of Biology at MIT in the summer of 2007 the scene was transformed. The days began slowly as graduate students and postdocs wandered into the laboratory, exchanging greetings but nothing more. The coffee machine was switched on. In half of the lab, benches had been replaced by long rows of desks each holding three or four computers. Each member of the lab gravitated to his or her own workstation and resumed his or her own project. Occasionally, one or another would leave their desks to refill their cup of coffee. Technicians and secretaries were nowhere to be seen. Few other physical objects ever moved in or out of the laboratory. Each desk sat adjacent to a small whiteboard that individuals used to make notes or write reminders to themselves. Although the lab housed several printers, they were only rarely used to print papers and diagrams, almost never data.

Here, most of the flow in and out of the laboratory was virtual; data, articles, results, diagrams, and ideas moved through the space invisibly. The most obvious architectural division is no longer between the bench space and the office space but between the ‘wet lab’ (full of lab benches) and the ‘dry lab’ (full of computers).

These differences in appearance and work indicate the fundamental changes that have taken place in biology in the last thirty years. Gilbert’s paradigm shift began to change the meaning of the very objects of biology itself. That is, computers have altered our understanding of ‘life.’ In the first place, this involved the ‘virtualization’ of biological work and biological objects: organisms and genes become codes of zeros and ones. But more importantly, information technologies require particular structure and representations of biological objects. These structures and representations have increasingly come to stand in for the objects themselves in biological work. Databases and algorithms determine what sorts of objects exist and the relationships between them. Compared to the 1960s and 70s, life *looks different* to biologists in the early 21st century.[[14]](#footnote-14)

The wet labs and wet work of biology has not disappeared, but it is increasingly dependent on hardware and software in intricate ways. ‘Seeing’ or analyzing a genome, to take one important example, requires automated sequencers, databases, and visualization software. The history told here is just not a story about how computers or robots have been substituted for human workers, or how information and data have replaced cells and test tubes in the laboratory. These things have occurred, but the changes in biology are far deeper than this. Nor is this just a story about how computers have speeded up or scaled up biology. Computers are implicated with more fundamental changes: changes in what biologists do, in how they work, in what they value, in what experimentation means, in what sort of objects biologists deal with, and in the kind of knowledge biology produces. ‘Bioinformatics’ is used here as a label to describe this increasing entanglement of biology with computers. By interrogating bioinformatic knowledge ‘in the making,’ we learn how biological knowledge is made and used through computers. These stories are not about the smoothness of digital flows, but about the rigidity of computers, networks, software, and databases.

Putting it another way, the chapters here all attempt to answer the question ‘What is bioinformatics?’ from slightly different perspectives. Bioinformatics does not have a straightforward definition; it overlaps in complicated ways with other terms such as ‘computational biology’ and ‘systems biology.’ It is a site of contestation about what biology will look like in the future: what forms of practice will it involve? How will knowledge be certified and authorized? What kinds of knowledge will count as biological? And in particular, what roles will computers and the digital play in biological knowledge? Bioinformatics entails competing answers to these questions. In this sense, bioinformatics is incomplete: it is a set of attempts to work out how to use computers to do and make biology. This is a pressing question precisely because computers bring with them commitments to practices and values of Big Science: money, large institutions, interdisciplinarity. Bioinformatics is not just about whether or not to use computers to do biology, but also about how to organize biological practice and biological knowledge more generally.

*Data biology and economies of production*

At the center of all these changes is data. Data that is the (virtual) stuff that is created by experiments, reduced in analysis, moved and arranged inside the machine, and exchanged and communicated through networks. Data may be numbers, pieces of text, letters of a DNA sequence, or parts of an image. When biologists are using computers, they are dealing with data in some way. This data is not knowledge. A biologist can submit data to a database, but he or she cannot submit it to a journal for publication.[[15]](#footnote-15) Its value lies in the fact that if it is appropriately treated and manipulated it can be used to create knowledge; but it is something other than knowledge.

Paying attention to data – rather than information – avoids suggesting that there is something inside the organismic body (code, information) that is somehow passed into the computer. Information is treated an immaterial message encoded in the chemistry of DNA that is then extracted into computer codes. Data, on the other hand, does not flow freely out from or across biological bodies – it is tightly coupled to the hardware and software within which it exists. Data belongs only to computers; it is part of a set of practices that only make sense with and through computers. Following the data helps us cut through the information metaphors and get closer to understanding of the perplexity of things going on inside the machine.

Data, and especially ‘big data,’ has received much recent attention from scientists. Biologists, along with astronomers, physicists, and Google’s engineers, have worried a great deal about how to extract knowledge from data. These concerns are centered not only on the need for new methods and techniques for dealing with large amounts of data, but also on the need for new institutions, new standards, new modes of work, new funding models, and new kinds of training in order to generate and support innovative methods of dealing with data. “Funding agencies have been slow to support data infrastructure,” Nature’s editorial reported in 2008, “…researchers need to be obliged to document and manage their data with as much professionalism as they devote to their experiments.”[[16]](#footnote-16) Finding ways to deal with data required new “incentives” for researchers.

Data is recognized as at once a technical and a social problem for science. As such, it provides a ready-made lens to understand how both scientific practice and scientific knowledge are being transformed through the production and treatment of data. Computers are – first and foremost – machines for the storage, processing, and analysis of data. ‘Following the data,’ then, provides a way of getting inside the machines, seeing how they are connected, seeing how things move around inside them, and understanding what constraints they impose. This means not only seeing how data is produced and used, but also what structures it inhabits and what kinds of practices and skills are required to maintain it.

Social scientists have, for some time now, been paying much attention to material culture. Historians of science in particular have shown convincingly how tools, objects, media, and so on have influenced the kind and character of scientific knowledge.[[17]](#footnote-17) The premise on which this investigation is based is that to understand bioinformatics, we need to supplement our attention to materials (and material culture) with attention to data (and data culture). Attention to material culture has showed us that power and knowledge is embedded not just in people and documents, but also in objects. Data draws objects into new relationships and new shapes. And these data structures, movements, and patterns contain knowledge and exert force on behavior. The ways in which data is generated, used, and stored constitute cultures that might be described and analyzed.

To understand computerized science and the virtual worlds that it creates, we must look to the practical imperatives of making, managing, manipulating, storing, and sharing data. Michael Fortun expresses the importance of data for contemporary science through his notion of ‘care of the data’:

It is through the care of the data that many genomicists work out and realize what they consider to be good science, and “the good scientist:” open to the data avalanche, welcoming its surprises and eagerly experimenting with them, but simultaneously cultivating a respect for the data flood’s openness to multiple interpretations, a keen awareness of its uncertainties, fragilities, and insufficiencies, a heightened sensitivity to its confounding excesses and its limitations.[[18]](#footnote-18)

However, what is at issue is not only that biologists (and other scientists) must be increasingly attentive to the subtleties of data management, but also that data increasingly exerts its own constraints on scientific knowledge-making. Data constrains what can and cannot be done in biology. This book follows the data to find out where biology is going.

The importance and ubiquity of data in biology is already obvious. As Fortun points out, genomics in particular is concerned with amassing more data, analyzing it in high-throughput, and storing it in ever-larger databases. This is a science already obsessed with data. The aim here is not to be swept up in this data flood, but to understand how data mediates between the real and the virtual and analyze its consequences for biological knowledge. It is not that life ‘becomes’ data, but rather that data brings the material and the virtual into new relations. Following the data provides a way of locating, describing, and analyzing those relations; it allows us to acknowledge that materiality is never fully erased, but that rather the material of the organism and its elements is substituted with other sorts of material: computer screens, electrons, flash memory, and so on.

Finally, paying attention to data can help us understand the economies of data production and exchange in biology. Just as Robert Kohler found that fruit flies generated a particular moral economy for *Drosophila* genetics, data biology produces particular economies of work and exchange. Following the data is not just about studying things inside computers, but also about showing how the flow and organization of data – between people, through laboratories – affects the organization and practice of biological work.

*Life out of sequence*

Describing the ongoing changes in biology demands a mixed set of sources and methods. In some cases, a properly historical approach is appropriate and in others more anthropological account is required. Scientific publications, unpublished archival sources, interviews with biologists, and observations drawn from immersive fieldwork in laboratories all find a place.[[19]](#footnote-19) The history of computers, biological databases, and algorithms are all a part of this story, but it is not a straightforward history of bioinformatics. There are detailed accounts of laboratory practice, but it is not just an ethnography. Rather, the combination of historical and ethnographic methods is used to answer historical questions about how and why biology has changed due to computing.

Six chapters chart the different paths of data into and through biology. The first two chapters examine the introduction of computers into biology as data-processing machines. Computers were developed as machines for data analysis and management in physics. This heritage is crucial for understanding how and why computers came to play the roles they did in biology. Thirty years ago, most biologists would not have given much thought to computers. The first chapter explains how and why this began to change in the 1980s. By 1990, a new sub-discipline had emerged to deal with increasing amounts of sequence data. Chapter Two characterizes this new field. Computers have re-oriented biology towards large-scale questions and statistical methods. This marks a break with older kinds of biological work that aimed at detailed characterization of a single and specific entity (for instance, one gene, or one signaling pathway).

The third and fourth chapters examine how data moves around in physical and virtual spaces. These motions constitute new modes of biological work. Drawing primarily on fieldwork in biological labs, these chapters describe how the exchange and movement of data means authorizing and valuing knowledge in new ways. Data-driven biology necessitates labs and work oriented towards productivity and efficiency.

The next two chapters examine data structures and how data becomes biological knowledge through databases and visual representations. Chapter Five details the history of biological databases, including the work of Margaret Dayhoff and the development of GenBank. It shows how databases act to structure and order biological knowledge. Chapter Six draws on published scientific literature, interviews, and fieldwork to piece together an account of the role of images as ways of making data into knowledge in computerized biology.

All this is by no means a chronological account of bioinformatics. It does not locate a definite origin, begin the narrative there, and follow the story through to the end.[[20]](#footnote-20) Instead, it presents several parallel accounts of bioinformatics, each with its own origins and unique trajectories. Bioinformatics is a manifestly interdisciplinary set of practices – it sits between and across biology, computer science, mathematics, statistics, software engineering, and other fields. Doing justice to these multiple and overlapping strands – networks, databases, software for manipulating sequences, and the hardware that they run on – means putting together an account that is ‘out of sequence.’ A combination of historical sources and ethnography (interviews and fieldwork) provide a series of cross-sections through the recent history of the biosciences. Each of these cross-sections shows, from a different perspective, how the flow of data has remade biological practice and knowledge.[[21]](#footnote-21)

What do all these perspectives add up to in the end? They show how biology is in the midst of a fundamental transformation due to the introduction of computing. It is not adequate to characterize this shift as ‘digitization’ or ‘informatization’ or ‘dematerialization’ or ‘liquification’ of life.[[22]](#footnote-22) Rather, we need to talk much more specifically about the ways that computers have changed what counts as interesting or answerable questions and satisfactory or validated solutions in biology. The material-to-virtual transition conceals a multiple and complex space of more subtle transitions: small to big, hypothesis to data-driven, individuals to teams, and specific to general. Following the data allows us to observe the effects of an increasing fixation on sequence. The particular ease with which sequences (DNA, RNA, protein) can be treated as data – represented, stored, compared, and subjected to statistical analysis by computers – has both transformed them as objects and rendered them increasingly central in biological practice.[[23]](#footnote-23) Our understanding of life, to an ever-greater extent, comes out of sequence.

But following the data also reveals ways in which sequence is arranged and manipulated. The ‘out of sequence’ of the title can also be read in the more familiar sense as meaning ‘out of order’ or ‘disordered.’ Bioinformatics reconstructs life not as a linear collection of genes, but as a densely connected network (of genes or proteins or organisms or species); the genome, we now see, is not in order; it is messy, mixed up. Data structures draw biological objects into new relationships, new topologies. Studying biology increasingly becomes about studying noise, randomness, and stochasticity. These are exactly the kinds of problems for which computers were first designed and used and which they are good at. Computers became widely used tools in biology only when these data-based modes of practice were imported from physics.

These three things – the ubiquity of computers and data in biology, the increasing importance of sequence, and stochastic approaches and understanding – are all linked. The use of computers in biology has provided epistemic space for particular sorts of problems – those based on large numbers, probability, and statistics. In so doing, it has also changed the objects that it studies – it has created new kinds of epistemic things. Sequences – and in particular the large sequence data sets from genomics – are particularly amenable to statistical analysis. It is the notion of the genome as a disordered object or an object ‘out of sequence’ that has allowed computers to become important tools for making sense of biology using statistical-computational tools. And, conversely, it is computers and computerization that have allowed sequences to continue to dominate the way we study and understand biology. The changing regimes of circulation and organization of biology are consequences of these epistemological shifts, rather than products of ‘informatization’ per se.

In the end, bioinformatics is likely to disappear, to evaporate as a relevant terms for describing biological practice. This will not happen because biologists stop using computers, but rather because the use of computers and bioinformatic methods will become so ubiquitous that it will make no sense to label it as a practice separate or distinct from biology generally. The kinds of knowledge and practices that are described here are already moving to the forefront of our understanding of life. If this turns out to be true, it is crucial that we reflect on the consequences of this change – what difference does it make that we now examine, dissect, manipulate and understand life with and through computers?

1. Dyson, “Our biotech future.” [↑](#footnote-ref-1)
2. A selection of readings: Forman, “Behind quantum electronics,” Kaiser, “Postwar suburbanization,” Kaiser, “Scientific manpower,” Galison, *Image and logic*. [↑](#footnote-ref-2)
3. Glaser, “The bubble chamber.” [↑](#footnote-ref-3)
4. Eisenhower, “Farewell.” [↑](#footnote-ref-4)
5. Weinberg, *Reflections*, 40. [↑](#footnote-ref-5)
6. Weinberg, “Criteria.” Elsewhere Weinberg spoke of the “triple diseases” of Big Science, “journalitis, moneyitis, and administratisis” (Weinberg, “Impact of Large Scale”). Another important critic of Big Science was the historian and physicist Derek J. de Solla Price – see Price, *Little science, big science*. A good historical summary of the debates is provided in Capshew and Rader, “Big science.” [↑](#footnote-ref-6)
7. Anderson, “More is different.” [↑](#footnote-ref-7)
8. At least in terms of funding, this ‘debate’ appeared to be quite one sided – with the high-energy physicists far out in front – until the end of the cold war; as Daniel Kevles has pointed out, the demise of the Superconducting Super-Collider left physicists in the 1990s searching for new ways to justify their work. See Kevles, “Big science and big politics.” [↑](#footnote-ref-8)
9. Latour and Woolgar, *Laboratory life*, 15ff. [↑](#footnote-ref-9)
10. Gilbert, “Towards a paradigm shift.” [↑](#footnote-ref-10)
11. This is partially based on the account of an early draft of this paper from 1990 that was examined by Fujimura and Fortun. See Fujimura and Fortun, “Constructing knowledge.” [↑](#footnote-ref-11)
12. Gilbert, “Towards a paradigm shift.” [↑](#footnote-ref-12)
13. Gilbert, “Towards a paradigm shift.” [↑](#footnote-ref-13)
14. Barnes and Dupre, *Genomes*. This view also is informed by the work on Brian Cantwell Smith on problems of ontology in computer science (Smith, *On the origin of objects*). [↑](#footnote-ref-14)
15. Although recently, recognition of the increasing value of data is changing this. For example, the journal *Gigascience* (launched in 2012 by the Beijing Genomics Institute) will accept and host large-scale life science data sets in its database. The aim remains, however, to promote and publish ‘studies’ of these data sets. [↑](#footnote-ref-15)
16. Editorial, “Community cleverness required.” This editorial appeared as part of a Nature special issue on ‘big data.’ [↑](#footnote-ref-16)
17. Some representative examples from sociology and history: Appadurai, *Social life of* things, Schlereth, *Material culture*, Bourdieu, *Outline*, Miller, *Material culture*. Some and representative examples from the history of science: Galison, *Image and logic*, Kohler, *Lords of the fly,* Chadarevian and Hopwood, *Models*, Rasmussen, *Picture control*, Rader, *Making mice*, Creager, *Life of a virus*, Turkle, *Evocative objects*. [↑](#footnote-ref-17)
18. Fortun, “Care of the data.” [↑](#footnote-ref-18)
19. During 2007 and 2008, I conducted participant observation-style fieldwork at three locations: the laboratory of Christopher Burge, in the Department of Biology at the Massachusetts Institute of Technology, Cambridge, Massachusetts; at the Edith and Eli Broad Institute, Cambridge, Massachusetts; and at the European Bioinformatics Institute, Hinxton, United Kingdom. [↑](#footnote-ref-19)
20. This is a good point at which to distinguish this account from Joe November’s very important historical account of the first uses of computers in biology (November, “Digitizing life” and November, *Biomedical computing*). November’s narrative begins with the first computers developed in the wake of World War II and concludes with the application of artificial intelligence to biology at Stanford in the 1960s and 70s. Here, I begin – for the most part – in the 1970s and attempt to come as close to the present as possible. In a sense, this account can be seen as a sequel to November’s: even at the end of the 1970s computers were extremely rare in biology and medicine, and those who used them remained far from the mainstream. To understand the ubiquitous, seemingly universal role that computers have come to play today we need to understand what happened with computers and biology *since* the 1970s. [↑](#footnote-ref-20)
21. Writing a history of bioinformatics is in some ways automatically an ahistorical project: no-one did ‘bioinformatics’ in the 1960s or 1970s. Of course, the sets of practices that became bioinformatics later on can be traced to practices from those earlier decades, but the ‘founding’ individuals were funded from different sources, organized in different disciplines, published in different journals, and understood their work as having significantly different aims. By tracing back these strands, I necessarily exclude or gloss over many of the other ways in which computers were used to do biological work from the 1960s onwards. Bioinformatics is centered almost entirely on a relatively narrow set of computation problems focused on DNA, RNA, and protein sequences. These have become by far the most visible ways computers are used in biology. But a computers were and are used for a range of other things (numerical simulations of population dynamics, simulation and prediction of protein folding, visualization of molecules, experimental design, simulation of evolution, simulation of the brain, instrument control, etc). This is not a history of all the ways in which computers have been used in biology. Rather the purpose is to understand why a certain set of problems and practices came to be picked out and to dominate our understanding of life. [↑](#footnote-ref-21)
22. Approaches that focus on ‘dematerialization’ in various forms include Mackenzie, “Bringing sequences to life,” Pottage, “Too much ownership,” Waldby, *Visible human project*, Parry, *Trading the genome*. Others (Lenoir, “Science and the academy” and Thacker, *Biomedia*), describe bioinformatics as a shift in the ‘media’ of biological work. [↑](#footnote-ref-22)
23. It is important to distinguish this argument here from arguments about gene-centrism that have been put forward by numerous scholars (notably including Keller, *Century of the gene*, and Nelkin and Lindee, *The DNA mystique*). Those arguments make the case that molecular biology has been dominated by a near-dogmatic obsession with genes and other associated regulatory elements, rather than epigenetic or cytoplasmic aspects of biology; DNA acts as a kind of ‘genetic program’ or ‘master molecule’ that controls all aspects of an organisms behavior. I agree with these assessments. However, my claim is slightly different in that it does not involve the claim that it is genes that are in control; rather, the sequence as a whole is the object of analysis. These could be RNA, DNA, or protein sequences, acting alone or in combination, with no molecule necessarily at the top of the hierarchy. All sequence, not just genes or DNA, contains meaning and importance for biological function which might be understood by performing statistical analyses. [↑](#footnote-ref-23)