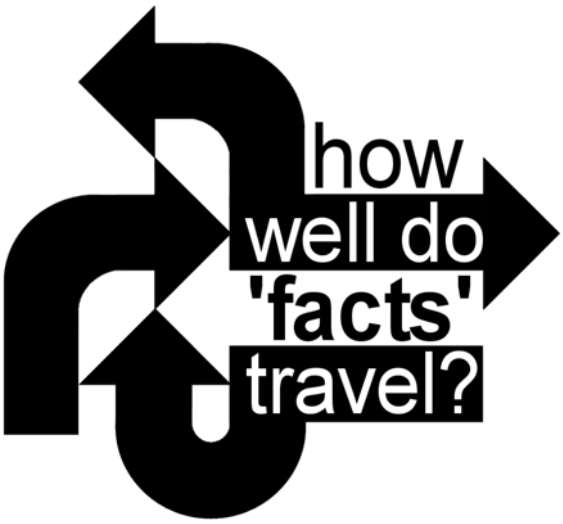
Working Papers on The Nature of Evidence:
How Well Do 'Facts' Travel?
No. 07/06

Wormy Logic: Model Organisms As Case-Based Reasoning

Rachel A. Ankeny

© Rachel A. Ankeny
Unit for History and Philosophy of Science
University of Sydney

March 2006



“The Nature of Evidence: How Well Do ‘Facts’ Travel?” is funded by The Leverhulme Trust and the ESRC at the Department of Economic History, London School of Economics.

For further details about this project and additional copies of this, and other papers in the series, go to:

http://www.lse.ac.uk/collection/economichistory/

Series Editor:

Dr. Jonathan Adams
Department of Economic History
London School of Economics
Houghton Street
London, WC2A 2AE

Tel: +44 (0) 20 7955 6727
Fax: +44 (0) 20 7955 7730

Wormy Logic: Model Organisms as Case-Based Reasoning[1]

Rachel A. Ankeny

Abstract

In the past few decades, so-called model organisms have become a cornerstone of research in the biomedical sciences. For the scientists, the model organism is both a practice ground for developing laboratory techniques, and a source of insights into common or even universal biological mechanisms. This paper examines the conceptualization of model organisms as models, and presents a formal account of how they are used to generate knowledge through what can be viewed as a form of case-based reasoning. Case-based reasoning is an epistemic process that is far from straightforward and may seem to fail to allow us to obtain the usual results we expect in science. Meanwhile, a growing literature within the history and philosophy of science on modelling and representation creates a space within which close attention to the principles and practices associated with such models may prove fruitful. Following a brief historical account of the development and use of one model organism, the nematode worm *Caenorhabditis elegans,* this paper addresses questions about the methodologies underlying work on genetic sequencing and developmental processes in this organism. In particular, what types of reasoning ground the use of experimental organisms when they are being developed and used as model organisms, and how are these models refined over time?

Introduction

> It's a motley collection of creatures: They fly, swim, wiggle, scurry, or just blow in the wind. But to the scientific community, this compilation has been elevated above all other species. They are the model organisms.[2]

Although various strains of numerous laboratory organisms have proven biologically and historically significant, "model organisms" have become a cornerstone of research in the biomedical sciences, especially

---

[1.] Written for *Science without Laws: Model Systems, Cases, Exemplary Narratives,* edited by Angela N.H. Creager, Elizabeth Lunbeck, and M. Norton Wise (Chapel Hill, NC: Duke University Press).

[2.] C. Bahls, J. Weitzman, and R. Gallagher, "Biology's Models," *The Scientist, 17 Supplement 1* (2003), 1.

in the past few decades. In addition to the mapping and sequencing of the human genome, among key components of the Human Genome Project (HGP) which officially began in 1990 was the mapping and sequencing of the genomes of non-human model organisms, including mice, nematode worms, flies, *E. coli,* and yeast.[3] James Watson has described the idea to include non-human model organisms in the HGP as his most important contribution to the project.[4] Despite this sort of support from early enthusiasts, some of the more contentious issues raised during the preliminary planning stages of the HGP related to the model organism projects, perhaps most importantly whether genetic sequencing was likely to result in knowledge that was relevant for the understanding and treatment of human disease processes, especially given the large amount of DNA without known function which was often derogatorily termed "junk DNA". Research on model organisms was rarely explicitly defended in the context of the project in its earliest days, perhaps in part because of assumptions about public and political perceptions and lack of ability (or desire) to understand this research, despite their explicit inclusion.[5] These organisms were used in the HGP as a means for developing the various mapping and sequencing technologies needed to study the more complex human genome, thus allowing these technologies to be tested and refined in a simpler, more efficient, and (purportedly) less expensive manner.[6]

[3.] See R. A. Ankeny, "Model Organisms as Models: Understanding the 'Lingua Franca' of the Human Genome Project," *Philosophy of Science, 68* (2001), S251-S261.

[4.] R. Lewin, "The Worm at the Heart of the Genome Project," *New Scientist, 127, 1731* (1990), 38-42; J. D. Watson, "The Human Genome Project: Past, Present, and Future," *Science 248* (1990), 44-48; *idem.,* "A Personal View of the Project," in D. J. Kevles and L. Hood (eds), *The Code of Codes: Scientific and Social Issues in the Human Genome Project* (Cambridge, Harvard University Press, 1992), pp. 164-173.

[5.] See discussion for instance in C. R. Cantor, "Orchestrating the Human Genome Project," *Science* 248 (1990), 49-51.

[6.] U. S. Department of Health and Human Services and U. S. Department of Energy, *Understanding our Genetic Inheritance. The U. S. Human Genome Project: The First Five Years FY 1991-1995* (Washington, DC, GPO, 1990). Note, however, that many model organism researchers participated in the HGP in large part to be able to study

But the genomes of these model organisms also were mapped and sequenced because they were expected to provide a basis for understanding normal gene regulation and human genetic disease, and more generally fundamental developmental, physiological, and other biological processes. Such expectations were based on the idea that many genetic and biological similarities exist between those organisms selected to serve as model organisms and humans; therefore model organisms would provide information that could aid in the interpretation of human genomic sequences and their products. This concept is rooted in the idea that there is conservation of many mechanisms and processes:

> Because all organisms are related through a common evolutionary tree, the study of one organism can provide valuable information about others. Much of the power of molecular genetics arises from the ability to isolate and understand genes from one species based on knowledge about related genes in another species. Comparisons between genomes that are distantly related provide insight into the universality of biologic mechanisms and identify experimental models for studying complex processes.[7]

Both the prevalence and centrality of model organisms in contemporary biomedical research, and claims about their use as the basis for deriving insight into certain common or even universal biological mechanisms, generate an ideal laboratory for examination of epistemic issues related to use of such organisms. In addition, the growing literature within the history and philosophy of science on conceptual issues associated with modelling and representation in science[8] and on

---

their organisms of choice in their own right, which in turn created various epistemic and pragmatic tensions within many laboratories and research programs, a point which I cannot examine in any detail here.

7. F. S. Collins, et al., "New Goals for the U.S. Human Genome Project: 1998-2003," *Science 282* (1998), 682-689, on 686-687.

8. For instance see M. S. Morgan and M. Morrison (eds), (1999) *Models as Mediators* (Cambridge, Cambridge University Press, 1999); S. de Chadarevian and N. Hopwood (eds), *Models: The Third Dimension of Science* (Stanford, Stanford University Press, 2004).

various model organisms[9] creates a space within which close attention to the principles and practices associated with such models may prove fruitful.

This essay examines the conceptualization of model organisms as models, and presents a formal account of how they are used to generate knowledge through what can be viewed as a form of case-based reasoning. Following a brief historical account of the development and use of one model organism, the nematode worm *Caenorhabditis elegans,* I address questions about the methodologies underlying work on genetic sequencing and developmental processes in this organism. In particular, what types of reasoning ground the use of experimental organisms when they are being developed and used as model organisms, and how are these models refined over time?

Some clarifications on terminology to begin: the term "model organism" is used throughout this essay rather than "model system" since the former expression is explicitly employed in the literature on the HGP and more generally in contemporary organism-based biology.[10] Model organisms can be seen as a specialized subset of the more general class of model systems, where the latter usually encompasses not only the organism but also the techniques and experimental methodologies

---

[9.] Though by no means an exhaustive list, historical and conceptual accounts of the development of and research with various model organisms that have influenced my research include A. E. Clarke and J. Fujimura, "What Tools? Which Jobs? Why Right?" in their *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences* (Princeton, Princeton University Press, 1992), pp. 3-44, and the essays in this collection; R. M. Burian, "How the Choice of Experimental Organism Matters: Epistemological Reflections on an Aspect of Biological Practice," *Journal of the History of Biology, 26* (1993), 351-367, and the articles contained in this special issue of the journal devoted to experimental organisms; R. E. Kohler, *Lords of the Fly: Drosophila Genetics and the Experimental Life* (Chicago, University of Chicago Press, 1994); A. N. H. Creager, *The Life of a Virus: Tobacco Mosaic Virus as an Experimental Model, 1930-1965* (Chicago, University of Chicago Press, 2002); K. A. Rader, *Making Mice: Standardizing Animals for American Biomedical Research, 1900-1955* (Princeton, Princeton University Press, 2004); and Ph.D. research in progress by S. Leonelli (Vrije Universiteit, Amsterdam) on Arabidopsis *thaliana.*

[10.] Compare J. A. Bolker, "Model Systems in Developmental Biology," *BioEssays, 17* (1995), 451-455.

surrounding the organism itself.[11] This essay explores some of the techniques and methods used to establish and refine model organisms, but primarily from the point of view, as it were, of the model organisms themselves.

Background: The Worm

C. elegans is a free-living nematode, around a millimetre in length, with extremely simple behaviours and structures, and a relatively recent history as a model organism.[12] As noted in the Nobel Prize for Physiology or Medicine presentation speech for 2002 which celebrated three worm workers and the "joy of worms", part of what makes it a good candidate for a model organism is that *C. elegans* is "loaded with features".[13] There are two sexual forms, a self-fertilizing hermaphrodite and a rarer male that can fertilize hermaphrodites, which differ slightly in appearance and structure; this feature makes it an excellent genetic system as organisms can either be pure bred by isolating hermaphrodites, or new genetic material introduced via breeding with males. The adult is composed of a tube made of an exterior cuticle, which contains two smaller tubes (the pharynx and the gut) and the

11. On model systems, see for instance H-J. Rheinberger, *Toward a History of Epistemic Things: Synthesizing Proteins in the Test-Tube* (Stanford, Stanford University Press, 1997).

12. For additional background on the history of the choice and use of *C. elegans,* see R. A. Ankeny, *The Conqueror Worm: An Historical and Philosophical Examination of the Use of the Nematode* C. elegans *as a Model Organism* (Ph.D. Thesis, Department of History and Philosophy of Science, University of Pittsburgh, 1997); S. de Chadarevian, "Of Worms and Programmes: *Caenorhabditis elegans* and the Study of Development," *Studies in the History and Philosophy of Science, 29* (1998), 81-105; R. A. Ankeny, "The Natural History of *C. elegans* Research," *Nature Reviews Genetics, 2* (2001), 474–478; K. F. Schaffner, "Genetic Explanation of Behavior: Of Worms, Flies, and Men," in D. Wasserman and R. Wachbroit (eds), *Genetics and Criminal Behavior* (Cambridge, Cambridge University Press, 2001), pp. 79–116.

13. Urban Lendahl for the Nobel Committee, Karolinska Institutet, Stockholm, Sweden, 10 December 2002, http://nobelprize.org/medicine/laureates/2002/ presentation-speech.html.

reproductive system. The organism is transparent throughout its life cycle, making observation of many biological processes possible by various forms of microscopy. The genome of *C. elegans* is approximately 100,000,000 base pairs, which is one-thirtieth the size of the human and twenty times that of *E. coli,* and was virtually completely sequenced as of December 1998.[14]

The choice of *C. elegans* by Sydney Brenner in the mid-1960s and the original pursuit of research focused on this organism primarily at a single institution (the Laboratory of Molecular Biology in Cambridge, England) to which most current-day researchers can trace their own lineages has resulted in a relatively cohesive community often celebrated as a model of scientific cooperation and shared understanding of fundamental concepts.[15] Hence, an analysis of how "the worm" (as it is called by researchers in this area and many in the broader scientific community) functions as a model organism can be used as the basis for understanding the epistemic structure underlying most ongoing research in this area.

General examination of the history of organism choice reveals that prospective model organisms typically are selected and constructed based not mainly on principles of or knowledge about the universality or even typicality of their biological characteristics and processes (though it is hoped that many features will prove to be shared or common to other organisms), but primarily due to perceived experimental manipulability and tractability. For example, *C. elegans* was chosen specifically for its developmental invariance and simplicity, despite the atypicality of these

---

14. *C. elegans* Sequencing Consortium, "Genome Sequence of the Nematode *C. elegans:* A Platform for Investigating Biology," *Science, 282,* 2012-2018; for general overviews of work on this organism, see W. B. Wood and the Community of *C. elegans* Researchers (eds), *The Nematode Caenorhabditis Elegans* (Cold Spring Harbor, Cold Spring Harbor Laboratory, 1988); D. L. Riddle, et al. (eds), *C. Elegans II* (Cold Spring Harbor, Cold Spring Harbor Laboratory Press, 1997).

15. L. Roberts, "The Worm Project," *Science, 248* (1990), 1310-1313; E. Pennisi, "Worming Secrets From the *C. elegans* Genome," *Science, 282* (1998), 1972-1974.

biological characteristics (among many others) of *C. elegans,* even in comparison to other closely related organisms. In short, the general aim of the original research project was to achieve an understanding of developmental processes in metazoans (animals with bodies composed of differentiated cells, as opposed to protozoa or unicellular animals), and in particular, the development of the nervous system, since it was thought to be the most complex and interconnected system in these organisms.[16]

Brenner wanted to do research with an organism which was experimentally straightforward to manipulate and had relatively basic behaviours and structures, but was not so simple as to be "unrepresentative". The goal was to "optimize" an organism, in large part through making a careful organismal choice to start, rather than focusing on achieving standardization once in the laboratory via inbreeding and other typical techniques. Brenner and most subsequent worm workers in the early years of the research implicitly assumed that although *C. elegans* is simple, it is similar to all (or most) of the more complex members of the metazoa in terms of the genetic control of cellular differentiation. In particular, the genetic control of the development of the structure of the nervous system was thought to be likely to have shared fundamental mechanisms, in large part because of an implicit assumption of genetic conservation, particularly of essential processes.[17]

One of primary ways in which *C. elegans* can be seen to be a model organism relies on the idea that a model has been established to which particular empirical instantiations (i.e., actual, material worms) can

---

16. S. Brenner, "Foreword," in W. B. Wood and Community of *C. elegans* Researchers (eds), *The Nematode Caenorhabditis elegans* (Cold Spring Harbor, Cold Spring Harbor Laboratory Press, 1988), pp. ix-x.

17. The idea of shared mechanisms can be taken to its extreme: Howard Gest has suggested that the literature surrounding the current proliferation of model systems (or organisms, to use my preferred terminology) often seems to use 'model' to signify universality, and has called for correction of what he considers to be linguistic mis-usage (which I would claim actually has much deeper, epistemic implications); see his "Arabidopsis to Zebrafish: A Commentary on 'Rosetta Stone' Model Systems in the Biological Sciences," *Perspectives in Biology and Medicine 37* (1995), 77-85.

be compared, in order to articulate variations and differences in various features. The use of this form of reasoning in is perhaps most familiar from basic genetics: the first step in the underlying strategy is to select and establish a "wild type" for the organism (taken as a standard from among other possible wild types available in nature) against which other genetic variants or abnormal types can be compared. Despite its name, the wild type may not be the most common, frequent, or even a "normal" version of the organism; sometimes it is simply the first strain that was discovered on which subsequent research has been based, but is oftentimes the easiest to manipulate experimentally. These experimental organisms of course are "natural", inasmuch as they are still actual, living, concrete organisms, and have been "selected from nature's very own workshop".[18] However, the carefully selected wild type is, in this sense, an idealized model of actual organisms in nature, since oftentimes they end up differing considerably from those highly rarefied beasts that remain isolated in the laboratory, particularly as a model organism comes to be more widely used.[19] Thus modelling occurs in most obviously in the establishment of the wild type, which is an essential first step to establishing and using something on an ongoing basis as a model organism. Without this, it is not possible to have a "norm" against which "abnormal" (or more precisely, that which is variant) can be compared, in terms of genetics, developmental lineages, and so on. So a worm that is abnormal in movement might be detected by comparison of the paths that it traces in response to a stimulus to those traced by a worm held to be "normal".

[18.] E. F. Keller, *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines* (Cambridge, Harvard University Press, 2002), p. 51.

[19.] In organisms where there is ongoing flow over time between the laboratory and the field or the wild, the amount of idealization in the model may be reduced, or more precisely, there may be more than one strain or variant that is held as a norm; however particularly with genetic model organisms (those selected primarily because of their power for genetic analysis, which is my focus in this essay), it is essential to settle on (and persist in using) one wild type.

A second way in which modelling occurs is in the establishment and use of what I have called elsewhere a "descriptive model".[20] The term "descriptive" is utilized to capture the idea that these sorts of models are descriptions which serve as prerequisites to explanatory questions; their articulation often is not motivated (at least immediately) by their future potential explanatory value. Thus in model organism work, there typically is an extensive research phase in which a descriptive model of the organism is developed. Consider, for example, the articulation of the "wiring diagram" of the neural connections within *C. elegans.* This model was a paper (and later computerized) series of drawings, which resemble electric circuitry diagrams.[21] The overall diagram was constructed by combining wiring diagrams from several individual wild type worms, not only because of practical or experimental limitations, but because it was deemed necessary to eliminate what seemed to be individual neural differences (even between genetically-identical organisms) in favour of a canonical nervous system. The wiring diagram is based on an abstract model of the worm in terms of the typical or usual neural connections exhibited not by any one specimen alone, or by numerous individual organisms, but by a more abstract construct hybridized from a few individual specimens. The wiring diagram thus is a model of the worm in terms of the typical or usual neural connections exhibited not by any one specimen taken by itself but by a very precisely derived type of construct.[22] This descriptive model is compared to the wiring diagrams

---

[20.] See R. A. Ankeny, "Fashioning Descriptive Models in Biology: Of Worms and Wiring Diagrams," *Philosophy of Science, 67* (2000), S260-S272. I am extremely grateful to Sabina Leonelli for her helpful critique of my overemphasis in this earlier article on the abstract features of *C. elegans* as model organism, which was due in part to my examination being solely about the construction of the worm's wiring diagram; I have attempted to clarify and remedy this narrow focus in the current essay.

[21.] J. G. White, et al., "The Structure of the Nervous System of the Nematode *Caenorhabditis elegans:* The Mind of a Worm," *Philosophical Transactions of the Royal Society of London: B. Biological Sciences, 314* (1986), 1-340.

[22.] This account has resonance with Jim Griesemer's analysis of material model building, inasmuch as the wiring diagram (as well as cell lineage and other descriptive

for worms that are variant or abnormal in neural patterns in order to assess possible connections between variations in genetic sequence and in neural structure, and eventually to test the range of the applicability of the descriptive model. Thus in this sense, some aspects of model organisms are in fact more like mechanical or physical models, constructed from natural organisms but constructed nonetheless, and hence highly idealized, since individual differences among wild type worms have been eliminated in lieu of (what are thought to be) the most commonly occurring structures.

Laboratory and community practices thus allowed the articulation and refinement of *C. elegans* as a model organism through at least two forms of idealization: the choice of a wild type (which provides concrete laboratory instantiations of the organism, permitting comparison for instance to particular mutant strains) and data-summarizing descriptive devices (such as wiring or cell lineage diagrams). Diverse model organisms undoubtedly have different histories that involve various kinds of modelling, depending on the natural features of the organism targeted to be exploited, the goals of the research community, and the degree of development of the model organism, among other factors.

The Principles of Case-Based Reasoning

The biomedical and human sciences have a long history of use of the case study as an object through which knowledge is generated and phenomena are made intelligible. The case is used to capture or summarize clinical and empirical data, to investigate underlying theories of disease, and communicate findings to other practitioners and

---

models associated with model organisms) can be seen as serving as "vicarious" models that serve as the basis for future theory development; see his "Material Models in Biology," in A. Fine, M. Forbes, and L. Wessels (eds.), *PSA 1990* (East Lansing, MI, Philosophy of Science Association, 1991), vol. 2, pp. 79-93.

researchers, among other purposes. To begin, it is helpful to provide a brief overview of the general form of case-based reasoning as used in medicine and elsewhere.[23] The basic method proceeds by construction of what might be termed (borrowing from the language of artificial intelligence) an "index case", in more or less detail depending on the goals of the situation in which it is to be used. In medicine, for instance, the index case often begins as a syndrome letter or published report on an individual patient, which then is abstracted into a model for more general use.

The case's use occurs through retrieval when a practitioner is presented with a new case which seems to have some overlap with the original index case, at least in terms of the details believed to be most relevant. This process involves a form of separating "signal" from "noise", to put it in different terms. The result is a feedback loop between processes of justification of the fit between the original index case and the new case under examination, particularly via assessment of similarity and identity relations. The outcome might not only be pragmatic (i.e., it may provide the basis for making a diagnosis or prognosis), but in addition, new cases can lead to modification of the index case as appropriate over time, or even adoption of a new index case for a particular condition, which in turn is disseminated through publication and teaching.

Underlying the index case and the feedback loop between it and any new instances, is an even more basic index case: that of the human being who is "normal" with respect to the abnormal features noted in the index case. What is considered to be the index case for the normal (i.e., the un-diseased condition) may also be altered over time, as the range of

---

[23.] This discussion summarizes a more detailed examination by me in "Case-Based Reasoning in the Biomedical and Human Sciences: Lessons from Model Organisms," in *Proceedings of the 2003 International Congress for Logic, Methodology and Philosophy of Science* (London, King's College Publications, 2005 forthcoming). On case-based reasoning in the human sciences, see especially J. Forrester, "If *p,* then What? Thinking in Cases," *History of the Human Sciences, 9* (1996), 1-25.

variants or errors in what was assumed to be the shared or common attributes (genetic, physiological, and otherwise) among healthy individuals are discovered. Thus the index case of the normal and of the disease condition often are constructed (and re-constructed) in terms of each other as more knowledge is gathered. What is essential in this form of reasoning is the feedback loop that exists between the descriptive model of the normal and the descriptive model of the abnormal condition. Newly-acquired evidence can change what is considered to be the index case or whether something should be considered to be a unique case at all.

Thus these cases are models inasmuch as, although they originate from some actual observed instance in the first place, once they begin to be disseminated and used, they become idealized away from particular details of the observed phenomena. They serve as intermediaries between the base of available knowledge (which is oftentimes overwhelmingly descriptive and relatively lacking in formal theoretical structure) and new natural phenomena which present themselves and which require understanding or explanation. Thus these types of cases can be viewed as serving as models, in the sense of Mary Morgan and Margaret Morrison (among others): they cannot be derived from either theory or data, and hence are partially independent. They clearly mediate between theory and the world, and are used in a “tool-like” manner to perform a range of tasks.

What is most important to note here is that as the index case is refined over time, a tension arises: in some sense, the base index case comes “closer” to what is really out there in nature, while at the same time it becomes more distant from any one concrete individual instantiation (any actual, material organism). Nonetheless, it remains a model, fulfilling many of the attributes that we expect from models: it is idealized, in that no patient typically fulfils all of the conditions captured in the

model, and yet patients can still be identified as having a condition or being an instance of that particular disease category or case.

Model Organism as Cases

The practices of contemporary biological science have (potentially conflicting) goals that are similar to those found in the practice of the medical sciences. There is a desire to get to the fundamental biological characteristics shared by all living things, be they biochemical, genetic, developmental, or neurobiological processes. At the same time, biologists are aware that any model system or organism selected for research may be problematic and atypical, particularly inasmuch as such systems are proving to be complex in ways previously that might not have been anticipated. The previous section on *C. elegans* as a model organism has shown several ways in which the organism as studied by biologists is an idealized entity or a model. The epistemic strategy of using the models as cases allows them to serve as a means of control of complexity, a way to create an appropriately simplistic yet descriptively rich basis for future studies and more traditional hypothesis testing, experimentation, and explanation.

Different aspects of a model organism thus can be viewed as index cases on which comparison to variant and abnormal instances of the same organism. So for instance, the wild type of the “natural” organism serves as an index case, in that it establishes a genotype which comes to be understood as “normal” and serves as the basis for comparison to subsequent cases of abnormal or variant genotypes. Similarly, the wiring diagram captures another sort of basic index case, to which variations in neural structure can be compared. Among the key foundational assumptions used to determine what counts as the relevant or most useful base index case for an organism are the anticipated degree of

genetic homology and genetic conservation. Eventually the index case may be altered to better reflect increased understanding of what is generalizable (or perhaps universal, at least within the species) in the model organism.

There are at least two important points implicit in this account that support the claim that the types of models discussed serve as the basis for case-based reasoning processes using model organisms. First, the idea that model organisms are in fact idealized entities as outlined has resonance with the construction of epistemologic entities elsewhere in the sciences, for instance of the "average man" in medical and human sciences, going back to the work of Quetelet: "The consideration of the average man is so important in the medical sciences that it is almost impossible to judge the state of an individual without comparing him to a fictive being that one regards as being the normal state and who is nothing but the [average man]".[24] Rich, descriptive idealizations thus are the starting point for case-based reasoning, as some base-line case must be provided to initiate the reasoning process. But these idealized cases are necessarily fictitious, as is the nervous system of the so-called "canonical worm", at the same time as they are essential tools for developing an understanding of the actual organism.

Second, note that as with medical case reports, usually there is no explicit (or implicit) testing of a hypothesis or theory, or what might be considered other "typical" scientific behaviours. Instead, the process proceeds by the proffering of observations and detailed descriptions, which may well point to testable hypotheses and explanations, particularly if they are to have an impact on the development of theory or on

[24.] A. Quetelet, *Sur l'homme et le developpement de ses facultes, ou Essai de Physique Sociale* (Paris, Bachelier, 1835), 2, 267, as quoted in J. Cole, "The Chaos of Particular Facts: Statistics, Medicine and the Social Body in Early 19th-century France," *History of the Human Sciences, 7* (1994), 1–27, on 12.

practice.[25] Thus there is a creation of an epistemological space or framework within which to ask questions. However, as bluntly stated by a commentator on medical reasoning, "with higher organisms, and especially with patients, it becomes hopeless to attempt to create complete descriptions…This is a kind of epistemologic surrender and consists in simply ignoring many of the things that could be truthfully said in order to say what must be said".[26] Both in medicine and in biological reasoning from model organisms, complexity, completeness, and perhaps "naturalness" are sacrificed in favour of selective construction of manageable material and a framework with which scientists can work and ask questions.

Case-based reasoning using descriptive models in biology thus relies on a double feedback loop between an index case and a case of interest that is abnormal or variant in some way. As in medicine, various cases are developed, for instance composed of descriptions of genetic or neural sequences in organisms. The base index case begins with a descriptive model of the organism that is established as being "normal" in phenotype, for which (say) the genomic sequence is identified and established as normal (or at least a norm against which other cases can be measured). This sequence can then be compared to that of organisms that are abnormal in phenotype (and thus assumed to be abnormal in genetic sequence) in order to draw out the functional properties of the genomic sequence within the particular model organism. Furthermore, an additional level of case-based reasoning occurs which then holds that determining the sequences in a variety of model organisms will reveal conserved (normal) genetic regions, which in turn will allow investigation of the same part of the sequence in the normal

---

[25.] See R. J. Simpson and T. R. Griggs, "Case Reports and Medical Progress," *Perspectives in Biology and Medicine, 28* (1985), 402–406.
[26.] M. S. Blois, "Medicine and the Nature of Vertical Reasoning," *New England Journal of Medicine, 318* (1988), 847–851, on 848.

human genome (or other, “higher”, organisms) and prove fruitful for understanding the functional properties of these sequences. Finally, the eventual goal is to understand the higher level, phenotypic results of abnormal, human genomic sequences found to be similar to the base case, namely the “abnormal” (or variant) sequences in the model organism, based on a correlation between these sequence and higher level properties such as disease conditions or other abnormalities.

Conclusions

> For even in the study of animals unattractive to the senses, the nature that fashioned them offers immeasurable pleasures in the same way to those who can learn the causes and are naturally lovers of wisdom… Therefore we must avoid a childish distaste for examining the less valued animals.[27]

What is most important to notice when analyzing the use of model organisms and particularly the way in which they function as a form of case-based reasoning is that answering the question of whether a model organism will in fact prove to be a useful model (i.e., for human genome sequencing) requires that researchers not only work on sequencing in the model organism but that this sequencing occur in tandem with sequencing in the object of interest, the human genome, and other comparative genomic work. This conclusion points to an important, but easily overlooked, aspect of modelling: in order for models to actually function well as models there must be ongoing refinement of the original descriptive models (the base index cases), as well as constant interplay between the original descriptive model and the subject being modelled (or the cases of interest or targets), and continuous development of the positive analogies between them (along with identification of the relevant

[27.] Aristotle, *De Partibus Animalium,* trans. D. M. Balme (Oxford, Oxford University Press, 1992), 645a8-17.

disanalogies and their import).[28] Much rhetoric surrounding model organism research unconstructively obscures this interplay and hence misrepresents the potential limitations of even good models. In other words, providing a model requires an interaction between the model and the object of interest being modelled, or between the base index case and the case of interest, including construction of similarity relations, which are impossible to devise without a detailed description of the process to be modelled (which in this case includes the functional properties of the sequence).

Case-based reasoning is an epistemic process that is far from straightforward and may seem to fail to allow us to obtain the usual results we expect in science, inasmuch as it fails (at least initially) to produce unified theories or mechanistic explanations, but instead results in a form of scientific understanding (perhaps of a weaker sort than our traditional theories and explanations) which is constantly evolving, incomplete, and uncertain, but nonetheless has the status of knowledge for its practitioners. Model organisms and the features of them that serve as cases mediate between theory and the world (and cannot be derived directly from either data or theories), and come to be used in a "tool-like" manner to perform a range of tasks, perhaps the most important of which is establishing a framework within which to ask questions.

[28.] Note the resemblance of this to views on models from the classic book by M. B. Hesse, *Models and Analogies in Science* (London, Sheed and Ward, 1963).

About the Author:

Rachel A. Ankeny is Senior Lecturer in the Unit for History and Philosophy of Science at the University of Sydney, of which she served as Director from 2000-2004. Her research interests range across the history and philosophy of the biomedical sciences and bioethics, and she has a forthcoming book focused on the development and use of the nematode *C. elegans* as an experimental and model organism.

Acknowledgements:

I am grateful to the organizers and participants at the original Princeton workshop at which an earlier version of this paper was presented for constructive feedback, particularly Angela Creager and Mary Morgan, and numerous other colleagues who have patiently supported and provided me with feedback on various incarnations of this research project. This research was partially conducted while a visiting fellow/faculty member at the Shelby Cullom Davis Center for Historical Studies, Princeton University, as well as at the Department of Economic History, London School of Economics, as part of the Leverhulme-ESRC sponsored project "How Well Do Facts Travel?" I also wish to thank my research assistant Fiona Mackenzie for sourcing literature.

LONDON SCHOOL OF ECONOMICS
DEPARTMENT OF ECONOMIC HISTORY

WORKING PAPERS IN: 'THE NATURE OF EVIDENCE: HOW WELL DO "FACTS" TRAVEL?'

For further copies of this, and to see other titles in the department's group of working paper series, visit our website at:
http://www.lse.ac.uk/collections/economichistory/

2005

01/05: Transferring Technical Knowledge and innovating in Europe, c.1200-c.1800
Stephan R. Epstein

02/05: A Dreadful Heritage: Interpreting Epidemic Disease at Eyam, 1666-2000
Patrick Wallis

03/05: Experimental Farming and Ricardo's Political Arithmetic of Distribution
Mary S. Morgan

04/05: Moral Facts and Scientific Fiction: 19th Century Theological Reactions to Darwinism in Germany
Bernhard Kleeberg

05/05: Interdisciplinarity "In the Making": Modelling Infectious Diseases
Erika Mattila

06/05: Market Disciplines in Victorian Britain
Paul Johnson

2006

07/06: Wormy Logic: Model Organisms as Case-Based Reasoning
Rachel A. Ankeny