Is Molecular Genetics Becoming Less Reductionistic?

Notes from recent case studies on mapping *C. elegans* and the discovery of microRNA

Preliminary version: Please request permission before quoting

Richard M. Burian

Virginia Tech

Blacksburg, VA

rmburian@vt.edu

Abstract

My answer to the lead question of the title of this talk is affirmative. But the point of the presentation is not simply to support this generic answer. Rather, I will explore some new findings bearing on how integrated networks regulate what genes 'do' in development – and how these understandings force geneticists to explain cellular and organismal traits and processes in ways that integrate environmental, cellular, and genetic causes. Put differently, given recent molecular findings, genetics must acknowledge diffuse causation of developmental processes and evolved phenotypes. (Part of the point: control of what genes 'do' involves parallel processing involving several distinct sorts of inputs, not just serial processing of genetic 'information'.) The result, I will argue, is a (weak) kind of (mechanistic) holism, incompatible with the strong genetic reductionism characteristic of the early days of molecular genetics. I will explore some methodological and epistemological consequences of these claims.

Introduction: The Shape of Our Project and Our Principal Thesis

Before I turn to the body of this talk, it will help to locate the details I will discuss within a much larger project. This is the first public presentation at the beginning of a project on which I have recently embarked with Kevin Elliott of the University of South Carolina and Maureen O'Malley, of the Egenis Centre at Exeter University. About three weeks ago, after extensive email discussions, we mapped out our initial plans in some detail and realized that we are already committed to at least four papers to put our argument together in full. Our general concern is to characterize a sea change that has taken place over the last 20 years or so in molecular biology, perhaps centering especially on molecular genetics. The title of today's paper hints at one central aspect of that change, which is also marked by the movement toward systems biology, post-genomic analysis of networks of interactions and the like. But the material has gotten so rich that I will only make it about half-way into what I had hoped to cover about microRNAs. I will try to preserve the main points I had hoped to make within the case study even though I will only touch briefly on the very rich developments after about 2002 in that very rich domain, a period that is critical for our main argument. We expect to argue that developments in this period demonstrate very nicely some key changes in molecular biology quite generally, changes that undercut the narrow molecular or genetic reductionism of earlier phases of molecular biology (particularly molecular genetics) without interfering with the mechanistic analysis of the details of molecular interactions. What will remain a major difficulty, probably for decades to come, is providing a clear account of the **modest holism** that we see emerging – a holism in which **what genes (and other determinants of cellular and organismal properties and behaviors) 'do' in a given cell or organism is co-determined by integrated networks of interacting molecules and cells where some of the causally effective properties of those networks and cells cannot be straightforwardly analyzed at lower levels.** That is, one or more integrated control networks are required for the formation and stability of cells and

organisms. Modest as it is, this holism recognizes a causal role for network, cellular, and organismal properties in determining various characteristic process of cells and organisms, properties that, we believe, cannot be fully determined by properties of the genes or the genome, or a full listing of the determinants, their spatial distributions, and their states at a given time. This provides a particular meaning for the claim that biological systems are open systems.

One way of thinking about this thesis is to put it in the context of the rise of genomics, systems biology, and high throughput technologies in current molecular biological disciplines. Where we used to be able to follow only one or a few molecules and processes at a time at a molecular level, we are now able to follow – literally – hundreds of molecules and at least a few dozen processes at once. This allows us to detect interactions among the entities studied and massive feedback and weak (or strong) integration of separable, modular processes and mechanisms in ways that were simply impossible before. It requires exploratory experimentation to establish the phenomena requiring explanation and, in terms of methods, new modes of interaction among exploratory experimentation, hypothesis construction, and hypothesis testing. This is the domain that we are exploring, working with case studies from several sciences and several interdisciplinary projects (such as the microbial tree of life, metagenomics (O'Malley 2007) miRNA research (Burian 2007) and epigenetics more generally, evolutionary developmental biology, and environmental toxicology (Elliott 2007)). My presentation today will reach the transitions we hope to explicate in detail, but I will have to do some handwaving rather than close exposition toward the end of talk because we are still very much at the beginning of our work. I am, for my part, all the more pleased to present this material at this early phase. I hope to get your feedback about how best to develop the line of work on which we have embarked and about errors or difficulties with which you think we must contend.

Before I turn to the body of my paper, I should state one sub-thesis, which involves a minority view of the character of a great deal of work in molecular biology and related sciences. This is the view that a great deal of molecular biology should not be viewed as hypothesis testing (though it has to employ hypothesis testing along the way), but as a kind of natural history – the natural history of biological molecules and the molecular process into which they enter. [Comment?: Note the difference between this point and the one Michel Morange made in a brand new paper (Morange 2008a), to the effect that at the origin of molecular biology many of the founders sought to 'naturalize' molecular biology by bringing it under physics. Michel's point is right – and the attempt to do just that was very influential, but it is entirely compatible with the sea change that we claim is taking, or has taken, place, to wit to reorganize molecular biology as a kind of natural history of molecules that depends heavily, as all molecular work must, on the tools and theories of the natural sciences.] Recognizing the importance and impact of this change helps to make better sense of what we are doing in our work on exploratory experimentation, which is already a major focus of the work that each of the three of us has been doing.

If you do not find the claim that molecular biology is a kind of natural history plausible, one way of generating intuitions along these lines is to consider what we do when we establish nucleotide sequences for preexisting genetic material or amino acid sequences for preexisting polypeptides. Such work is enormously experimental and interventive, but it has gone wrong if it does not yield the exact sequence(s) of the molecule(s) of concern *as they were before the intervention took place*. Similar things can be said for the much more complex sequences of molecular events involved in development when, say, *C. elegans* (or any other developing organism) proceeds from one developmental stage to another. Which molecules are where? Which molecules did they interact with? How are they altered in various circumstances? These questions are natural historical questions that require solution.

Many biological mechanisms and processes are studied in this natural historical mode. Transcription of genes, splicing of primary RNAs, translation of mRNAs into sequences of amino acids are nice examples. It is only when the classificatory and descriptive work required to understand these processes and mechanisms has made substantial headway, and the molecules involved in various circumstances have been at least partly identified and partly characterized, that we are able to produce detailed theoretical analyses of fundamental mechanisms and processes at the molecular level. Lest you think I am being genocentric here, similar things can be said about photosynthesis and metabolism, and about proceeding from one developmental stage to another in the history of an organism – indeed, for virtually all major cellular and organismal processes.

Today, I will report sketchily on two emblematic case studies that nicely illustrate many of the claims I have just outlined. The first concern Sydney Brenner's and his colleagues' mapping of C. elegans, the second concerns the discovery and subsequent exploration of the functions and roles of microRNAs. After a considerable amount of exposition, I will return to a general discussion of what recent work on microRNA is about, and to the general issues I have raised in this introduction.

Skip to the Central Dogma

Our project is, obviously, still in a formative stage, so I shall not set the background for the entire project, but, rather, sketch a synopsis of my own path toward our position.

- Long-term focus: **the interactions among developmental biology, evolutionary biology and genetics**. Anyone who has paid attention to the history of pre-molecular genetics should know that Mendelian genetics had nothing like a sufficient explanation for developmental processes or phenotypes and that there were deep and

fundamental conceptual disparities among workers whose primary disciplines were embryology (or developmental biology), evolution in the tradition of the synthetic theory or population genetics, and geneticists. It is of great interest to see how these conceptual disparities, which used to appear intractable, are being resolved, slowly and with difficulty, in the new molecular forms of the relevant disciplines and/or their re-formed molecular successors.

- A second long-term focus: **exploratory experimentation in biology**, especially its roles in molecular biology. This is how Kevin and Maureen and I formed our collaboration. At last summer's meeting of the International Society for History, Philosophy, and Social Studies of Biology, we gave three wholly independent papers on exploratory experimentation in molecular biology and, to our surprise, found out that our interests and views matched exceptionally well. As most of you already know, what might be called broad exploration, the sort of thing that can be accomplished with the high throughput technologies of contemporary molecular biology, has played a major role in the transformation of a large number of disciplines.
- One of the relatively strong views about molecular biology at which I arrived over the last ten or fifteen years, now found in a number of places and supported in various ways by various authors: A substantial part of molecular biology is natural history and, as such, is largely descriptive, not fundamentally based on hypothesis testing. Such natural history, however, is strongly experimental, for it is the natural history of interacting biomolecules, which cannot be accomplished without reliance on difficult, often highly interventive and arduous experimental protocols (which, in their own right, often involve considerable amounts of hypothesis testing along the way, used in the effort to establish the reliability of the outcomes and interpretations of the

experiments). This is the first step toward appreciating the necessity of working out a methodology in which hypothesis testing and exploratory experimentation are closely, perhaps inextricably entwined.

- In thinking about the previous two points, Kevin, Maureen, and I recognized that part of our project must be to understand how the methodologies of natural historical description and classification and exploratory exploration can and should be integrated with standard accounts of hypothesis testing. The point here is partly practical and partly normative: granting agencies, by and large require that one couch one's proposal in terms of the hypotheses to be tested, which does not fit well with the practice of molecular biology when it is doing natural history or exploratory experimentation. And the most exciting work of all is in the rebuilding of theories and hypotheses, and reorganizing and restructuring hypothesis testing procedures in light of novel findings about useful, well grounded classifications of molecules and molecular processes. We think that the contrast between exploratory experimentation and hypothesis testing should be blurred in the new genomics and systems biology and that the current emphasis on hypothesis testing in awarding grants and in methodology and philosophy of science tend to undervalue and hamper important lines of molecular work.

If you do not find this plausible, one way of generating intuitions along these lines, is to consider what we do when we establish nucleotide sequences for genetic material or sequences of amino acids for polypeptides. Such work is enormously experimental and interventive, but it has gone wrong if it does not yield the sequence of the molecule(s) of concern *as it was before the intervention took place*. Similar things can be said for the much more complex sequences of molecular events involved in development when, say, *C. elegans* (or any other developing organism)

proceeds from one developmental stage to another. Which molecules are where? Which molecules they interact with?, how they are altered in various circumstances? These questions are natural historical questions that require solution. Many biological mechanisms and processes are studied in this natural historical mode. Transcription of genes, splicing of primary RNAs, translation of mRNAs into sequences of amino acids are nice examples. It is only when the classificatory and descriptive work required to understand these processes and mechanisms has made substantial headway, and the molecules involved in various circumstances have been identified and at least partly characterized, that we are able to produce detailed theoretical analyses of fundamental mechanisms and processes at the molecular level. Lest you think I am being genocentric here, similar things can be said about for photosynthesis, metabolism, and for proceeding from one developmental stage to another in the history of an organism – indeed, for virtually all major cellular and organismal processes.

- My colleagues and I hold that the rise of genomics and systems biology can be used to help characterize the fundamental changes that have come about in several domains of molecular genetics and molecular biology during the last twenty years or so. I hesitate to provide a crude shorthand description of what we think is at the heart of the change, for there are enormously many complex changes and issues involved here. Nonetheless, it will help, I think, to put the matter as straightforwardly as possible, so I will provide a shorthand marker, leaving discussion of qualifications and amplifications to the discussion period. Thinking about the matter from the perspective of genetics, which is where I got started on these issues, **genetics has been hoist on its own petard**. Put more politely, insofar as genetics thought that fundamental properties of organisms are determined by the genes,– a claim which

many geneticist have made and supported strongly – **genetics itself has demonstrated the insufficiency of genes or genomes, by themselves, to determine most of the phenotypes of concern**.

- Over the last three years I have begun to study **epigenetics and miRNAs** seriously. My interest was sparked in part by the roles played by exploratory experimentation in the experimental work involved and in part by the ways in which these lines of work make salient the important sea change I have been describing. Let me illustrate the change in terms of the attitude, even at the center of molecular genetics, regarding **the central dogma**. In eukaryotes, if one examines the issues affecting **which polypeptides are formed from a gene, which proteins are made from a given polypeptide (i.e. how a particular string of amino acids derived from a particular sequence of nucleotides in a mRNA is formed into distinct functional proteins in different cells), and the processes that allow variants of specific genes to trigger specific developmental change**s, the answers always require **the input of a network of cellular regulatory controls that are, at best, indirectly genetically determined**, with the intermediation of numerous non-genetic factors, including the history of the cell lineages in question and the influence of exogenous signals exogenous to the cell. This is the entry point for *C. elegans* into this paper. The work of Sydney Brenner, the C. elegans workers around him, and the people in the intellectual lineage descended from him provided core breakthroughs for making these points. The work on *C. elegans*, which proved to be an ideal model organism for the purpose, and the decisive articulation of many of the relevant points by Brenner and other *C. elegans* workers, were of considerable historical importance in influencing the changes in which we are interested.

There. Now you should have a sense of the points toward which we are working and which will be illustrated in today's lecture. As I proceed, you will see how the particular cases we have chosen help us to articulate our account of the major changes that we will examine in the larger project. We will, of course, address many other cases and disciplines in the project as it develops and have already started on projects on the microbial tree of life, genomics, systems biology, nanotoxicology, and funding philosophies in granting agencies. Although these other projects are not yet far enough advanced to be comfortably integrated with the work on which I am reporting today, the two cases I will outline are emblematic and nicely illustrate many of the claims I have just outlined. It is time to turn to Sydney Brenner and his colleagues' mapping of C. elegans the discovery of microRNAs. After a considerable amount of exposition, I will return to a general discussion of what recent work on microRNA is about, and to the general issues I have raised in this introduction.

Resume Here

The Central Dogma

To get a better handle on how genetics and molecular biology have changed in the last couple of decades, it is useful to consider the standard interpretations of the central dogma, especially in its hardened versions. As you all know, the Central Dogma was first publicly enunciated together with the sequence hypothesis in a famous paper by Francis Crick, "On protein synthesis" (Crick 1958). In Crick's version, information – by which he meant mainly sequence-determining information – could move from nucleic acids to proteins, but was not vice versa. Put differently, nucleic acid sequence information was used, biologically, to determine amino acid sequences, but amino acid sequences were not and, indeed, probably

could not be used, biologically, to set the sequences of nucleotides in an organism. The sequence hypothesis stated that there is a specific mapping from nucleotide *sequence* that determines amino acid sequence. (Notice the connection of these interlocked doctrines to the Weismannian doctrine that heredity is restricted to what can be transmitted through the germ line, applied to the discovery that DNA is the genetic material and the view that the sequence of nucleotides is the salient variable in the structure of the DNA molecule.)

Crick did not originally mean the dogma to be taken on faith; it was a dogma because it was *unproven*. It should be employed heuristically, but, ultimately, research needed to go behind it to test or revise it (Crick 1970, Olby 1975). However, many key figures, e.g., Jacques Monod (Crick 1970, Olby 1975) took it to be a dogma of the other sort, i.e., to be taken on faith. And James Watson, propagated a stronger, linear version of the dogma, in which information covered much more than sequence, and the flow of information was unidirectional: once it left DNA, it could not go back. This diagram comes from Watson's extraordinarily influential *Molecular Biology of the Gene*. Most early molecular geneticists employed his model and (largely unconsciously, I think) set aside the other molecular properties of DNA and proteins (e.g., their three dimensional conformations and temporal changes in their conformations) when they were thinking about hereditary information. The dogma then meant that, unless something untoward intervenes, the genes determine the structure and conformation of proteins in the appropriate cellular contexts, and largely determine the sequence of events and cellular processes that go into the development of an organism. And the genes provide sufficient information to yield the fundamental explanation of 'good' phenotypes, e.g., those that show Mendelian regularities or that can be produced by specific somatic cell mutations (as in the development of cancers). [For a debate bearing on these claims, see (Burian 2009, Sapienza 2009).]

As molecular biology 'hardened' into a discipline (or something like a discipline), the Central Dogma, interpreted along the lines advocated by Watson, acted as an article of faith and played a strong role in shaping mainstream thinking, even when the sequence hypothesis acquired a much more nuanced interpretation thanks to such complications as split genes and alternative splicing. Recently, partly thanks to the discovery of gene sharing, of mechanisms of epigenetic inheritance, and of the enormous role of regulatory networks in controlling what is made out of any given stretch of genetic material, interpretations of the central dogma have softened. Not coincidentally, several historians and scientists have reviewed the history of the central dogma, among them Bruno Strasser and Michel Morange (see Morange 2006, Morange 2008b, Strasser 2006) and have shown that the story of the last 50 years is more interesting than the stereotyped capsule history I just recited suggests. The work initiated by Sydney Brenner on *Caenorhabditis elegans*, and some of the subsequent developments in the use of C. elegans as a model organism nicely illustrates some of the complexities involved. *C. elegans* research is an excellent marker of the changes crucial for the larger story I sketched in the introduction to this talk. So it is time to turn to *C. elegans.*

A Sketch of its Early Evolution of Brenner's Worm Project

The nematode *Caenorhabditis elegans* is a major model organism (Ankeny 1997, 2000, 2001, 2007, de Chadarevian 1998). It has contributed greatly to the shaping of the Human Genome Project, thanks in good part to the extensive and detailed maps of cell lineages, of neural and other structures, the linkage and physical maps of genes, the mapping of nucleotide and amino acid sequences and of sequences of interactions in molecular process built up under the aegis of a project initiated by Sydney Brenner in 1964 or 1965. The *C. elegans* sequence was the first animal sequence completed in connection with the HGP; its sequence was completed as a model ahead of all the others and a leader of the *C. elegans*

sequencing project, John Sulston was put in charge of Great Britain's contribution to the HGP (de Chadarevian 2004). Although the organism has long been known to biologists (early work on it was done in France (Maupas 1900)), its role as a model organism can be traced back directly to Sydney Brenner.

In 1962, Brenner a close collaborator of Francis Crick's in Cambridge, started a systematic search for a small metazoan to use in studying behavior and development and their genetic regulation, and the structure of a simple nervous system to understand the influence of that structure on simple behaviors. In advance of making his choice, Brenner set out a series of criteria that such an organism should fulfill. Some of those criteria are hinted at on this slide, which is part of a 1963 list of the advantages of *C. briggsae*, which ultimately was second choice after *C. elegans*. By 1965, he had started the project with a couple of technicians, slowly adding a few people, one or two at a time at first, assembling a small and then a middle-sized group of remarkable individuals who shared his intense sense that the work must proceed by filling in all of the details, making a series of maps that ultimately interlocked with each other to provide the entire developmental history of the organism. His choice of C. elegans was thought out very carefully; it was a wise choice given his aims, but it was also aided by great good deal of luck along the way. *C. elegans* has the most stereotyped developmental known for any organism of comparable size. The N2 strain from Bristol, chosen as the wild type, has significantly better handling properties and certain regularities not found in other strains. Its features have dictated larger understanding of what is standard across a much wider range of organisms (though not all features thought to be standard really are so) (Ankeny 2007, Félix 2008). In any case, by intensely concentrated grunt work at the very limits of available technology, the entire cell lineage of the wild type was mapped in about a decade, first for some parts of the nervous system and then for the entire organism except for the variable cell divisions in the germ line, linkage and physical

maps of the genome were constructed, and the effects of a number of mutations traced and mapped in considerable detail.

The spirit in which this was done emanated from Brenner, and it characterized much of the work done by the first three or four generations of people conducting research on 'the worm', their label for *C. elegans*. The spirit is that of descriptive natural history, carried out at the cellular and molecular level by building what Rachel Ankeny has called 'descriptive models' of developmental and physiological processes, cell lineages, genetic and gene regulatory networks, and all the relevant structures of the worm (Ankeny 2001). Most of the early questions were descriptive, and most of the initial descriptions required hard investigative work to fill in obvious lacunae in what was known. Consider, for example, the determination of which cells produce which other cells in a cell lineage? Each cell division must be followed in detail, and the similarity in different worms of the same strain of the sequence of cell divisions and cell deaths must be vouchsafed. Fortunately, the stereotyped nature of *C. elegans* development means that the cell lineage is the same for effectively all wild-type N2 worms under a wide range of moderately standardized cultural conditions. And the deviations in a surprisingly wide range of circumstances, that is, with specific sorts of shocks at specific times or specific mutations, are equally stereotyped. Rachel Ankeny has developed a valuable account of descriptive modeling, as developed by the worm researchers, and the results it produced (Ankeny 2000, 2001, 2007). Her account raises a number of important methodological and philosophical questions, but I can't stop to pursue them further today.

Brenner sought to find out how genes control relevant traits, very much in the spirit of the central dogma. He was thoroughly convinced that there is, in some appropriate sense, a genetic program that controls development and many specific behaviors. He sought to ascertain, in molecular detail, how such control was exercised in a simple metazoan with

moderately complex development and complex behaviors, focusing in part on whether and how the one-dimensional sequence of nucleotides might possibly specify the developmental sequence of extremely elaborate and complex three-dimensional structures – i.e., how to get from a one-dimensional input to a result that was at least four dimensional (and probably much higher). This meant pursuing the mapping projects in excruciating detail to characterize what happens and how changes in one place alter what happens later, not only at the cellular level, but also at the molecular level where possible. At both the cellular and the molecular levels, such mapping was far beyond the capability of the technologies available in 1965, so a great deal of time, effort, and tedious work went into developing techniques and trying new ways to answer simple questions in great detail (what?, when?, where?, who interacts with whom?), eventually down to the level of macromolecular sequences and structures. This phase of exploratory experimentation to find techniques to answer these 'simple' questions and to answer them without particular hypotheses regarding the content of the specific answers is of considerable interest for our larger project. For today, it suffices to say that such an effort not fit comfortably with standard accounts of hypothesis testing and that it demanded enormous ingenuity and an amount of patience, plus much difficult experimental work before it produced any significant results.

This minimal description of the beginning of the worm project suffices for present purposes. It was nine years before the first full papers were published (Brenner 1974, Sulston and Brenner 1974). But once the worm work took off, it grew quickly within the bounds of a very close community; all of the first and second generation worm workers and most of the third generation spent at least some time in Brenner's group at the Laboratory of Molecular Biology in Cambridge or were in close contact with that group. Results were shared immediately (as the work became intercontinental, the sharing developed electronically along with DARPANET). Descriptive mapping remained focal for at least four generations of

researchers. In every map, as details were acquired, specific lacunae were uncovered. Efforts to fill in details of particular maps often involved generating and testing 'local' hypotheses relevant to the issue at hand (e.g., about the stage at which the fate of a cell's daughter was determined and whether the fate of that daughter depended on subsequent interactions with other cells). We will follow an example of this sort concerning molecular interactions when we turn to the discovery of microRNAs in a few moments. But before doing so, it will help to spell out a couple of lessons from this sketch of the early stages of the worm project, lessons that can be drawn from even this minimal account of an extraordinarily rich history.

As was true for Darwin in building his theory of descent with modification, the first major job of the worm project was to build up detailed knowledge of extensive domains and patterns of phenomena potentially relevant to the core problems of interest. Just as Darwin had to master the phenomena of biogeography, breeding, embryology, paleontology, systematics, and variation among related organisms before he could determine exactly what explananda required detailed explanation, and just as Darwin's explanations and the criteria for evaluating alternative explanations were shaped by his knowledge of the phenomena, so the worm workers had to establish the genetic content, the timing of gene activation, the genetic linkages and cellular lineages, the structure and connections of the nervous system (including the wiring diagram of all neural connections), the effects of mutations on particular cells or lineages, the correlation between nervous system structures and behaviors, and alterations of behavior in the presence of changes in any of the structures, connections, mutations, etc. before they could proceed to develop and test theories and hypotheses at the level in which they were interested. In philosophical terms, reorganizing and elaborating the explananda, that is, determining what precisely it is that calls for explanation, is the first step

in devising better explanations and testing them. Darwin and the *C. elegans* workers illustrate the virtue of this way of working. Both started from fairly vague starting points, they made moderately precise what phenomena had to be accounted for and how the patterns of phenomena fit together in the very process of providing an orderly account of the phenomena. Only with such information in hand did they construct and test serious explanatory hypotheses about how the phenomena, processes, and patterns of interest were controlled, regulated, or brought about.

Brenner's general views about how best to characterize the developmental phenomena and processes, and the inputs governing behaviors of the worm evolved considerably during the long exploratory period as did his account of the genetic determination of cellular structures and lineages, and of development and behavior. The changes stemmed from what the mapping projects showed, for example, about the timing, localization, and controlling steps in forming cell lineages and the responses of the cell lineages to various contingencies at different times during ontogeny. Such developmental phenomena, the ways and pathways by which mutations or changes in the structure of the nervous system affected various behaviors, and also many other related phenomena had to be taken into account. Thus his views on genetic determination were backed up by an enormously detailed account of the phenomena to be explained, the sequences of events, the molecular and cellular players that influenced those events and the timing with which various switches were thrown. Although he never abandoned the idea of a genetic program underlying development and the role of that program in determining even some molar behaviors of complex organisms, his account of the ways in which such a program achieved its effects departed quite significantly from strong readings of the Central Dogma. I take the liberty of quoting a substantial part of two paragraphs from a 1998 paper at some length to illustrate the direction in which he moved and how it was connected to this style of work. I hasten to add that there are texts that show

him to be working in this direction as early as 1973, the year before the first major publications of the worm project appeared. (These were his paper on the genetics of *C. elegans* and Sulston and Brenner's paper on the DNA of *C. elegans* (Brenner 1974, Sulston and Brenner 1974)). (For some early texts illustrating Brenner's view see the Appendix.) Here then is Brenner's evolved statement of the position he started taking in 1973, as of 1998:

> … [L]iving systems… are totally unlike all other natural complex systems, in that they carry an internal description of themselves written in their genes. It is this description which is passed on from generation to generation and from which the organism is 'computed'. If we compare this to the weather, for example, we find that there is no internal description of the weather that we can separate physically from the weather itself. For the weather we need the physics of matter and energy, but the existence of DNA implies something new; it is the physics of information, that is, computation.
>
> … [Schrödinger] was clear that the genetic material contained a programme for the development of the organism, but he thought that the genes also contained the means for its execution. They do not contain the means, but, rather, a *description* of the means for execution. This was precisely the distinction made by John von Neumann in his theory of self-reproducing machines, and shown to be a necessary feature of such automata. The means to translate the instruction tape is obtained from the parent machine and is used to read the description of the means and so install the means in the daughter machine. In biological systems, the egg has the means to read the genes, and the new organism makes new eggs. Thus, in addition to DNA, there is a physical continuity of the reading machinery over the total course of biological evolution, but the informational continuity is preserved in the genes. [p. 107]

Brenner S., 1998, 'Biological computation + Discussion'. In Bock G.R. and Goode, J.A. (eds.), *The Limits of Reductionism in Biology*, Chichester: John Wiley, 106-111 + 111-116.

The evolution of Brenner's views and their exact interpretation raise significant difficulties; they are worth a separate paper. But these quotations provide a good sense of the direction in which he moved. On one reading, which I favor, but which probably pushes beyond his ultimate position, the genetic description of the organism encoded in the genome, encoded in the one-dimensional string of nucleotides, is not by itself sufficient to produce an organism because there is additional heritable information in the structures and systems that interpret genetic information. In many ways the *action* is in the machinery that does the reading out of the otherwise-inert instructions. Indeed, since different cells read the same instructions in different ways, with different results (cf., for example, alternative splicing and the many allied phenomena), and since the regulatory apparatus must be set to do different jobs in different cells, what a gene or a long string of nucleotides 'does' in development is determined not by the gene itself or the genome, but by how it is read out and how its immediate products are handled in various contexts. The dimensionality of the linear string of nucleic acids is not sufficient to yield instructions to produce an organism except in the context of a properly set-up cell or system of cells with structured molecular contents which operate on not just on that string of nucleotides but also on the other ordered materials contained in the cell itself. Only then can the cell interpret the DNA as containing *these* instructions rather than *those*. Add that there are heritable changes in the structure of the machinery not controlled by the DNA and you get a pathway leading to epigenetics and to some of the many new discoveries about various regulatory systems and networks of regulation.

I cannot amplify on this reading or follow Brenner's views further, at least not if I am to reach the other, related, case study that will take us straight to the transition from the older molecular biology to the world of genomics and systems biology. This is the work done on C. elegans leading to the discovery of microRNAs.

[Stop to ask: How many in the audience have some sense of what microRNAs are and what sorts of jobs they do.]

The Discovery and Exploration of miRNA

The first point to make about the discovery of microRNA is that it was accomplished in the course of very difficult work within the more-or-less 'normal science' program of mapping the worm that I have been describing. Very difficult, because it was unquestionably pushing the limits of available technological tools. And the idea that the work was 'normal science' must be taken with a grain of salt. But it probably felt roughly like 'normal science' to the investigators at the heart of the process, for they were filling in some curious and quirky lacunae in the gene regulation map. Their results, in the end, forced major changes in our understanding of the mechanisms of gene regulation, but the story begins as a straightforward attempt to fill the lacunae in the network of regulatory interactions affecting expression of particular genes. I will present only a capsule summary here; the details are enormously richer than can be conveyed in brief compass, but the outline of the story is fairly straightforward.

One line of work in constructing the lineage maps was to understand the switches that are thrown in order to alter cells or produce variant cells in the transition from the embryo to each of the four larval stages and the adult stage. With long hard work, about 90 so-called *lin* mutants (i.e., mutants that altered cell lineages) were characterized and their effects tracked.

About 120 *let* (for lethal) mutants also proved to be of interest for this purpose, for some of them interacted in the same regulatory networks as some of the *lin* mutants. These regulatory networks, in line with the standard account at the time, were (so far as was then known) composed of proteins, themselves, of course, under genetic control, and of the genes themselves plus the messenger RNAs that carried information extracted from those genes. This so-called protein orthodoxy – that the principal regulatory molecules are proteins acting on mRNAs and DNA – benefitted from and reinforced the Central Dogma. The protein orthodoxy limited the tools used to understand gene regulation (and hence, ultimately, the genetics of development) to ways in which the timing of gene expression was regulated indirectly by the genes themselves together with their more-or-less direct products, mRNAs and proteins.

For the discovery of miRNA, the stories of two genes are of special interest. [I am borrowing heavily here from reviews by Eric Lai and Gary Ruvkun et al. (Lai 2003, Ruvkun et al. 2004) as well as a reading of the original sources.] Much of the relevant work on these was carried out in two labs, those of Victor Ambros and Gary Ruvkun, though, of course, other labs were involved at every stage of the work. For a long time, the work of interest centered on one of the so-called heterochronic mutants, *lin-4* and its interactions with another heterochronic mutant, *lin-14* as well as related genes.. [Heterochronic mutants are those that alter the timing of switches of cell type, block such a switch from acting, or alter the timing of specific developmental processes.] By the mid-1980s it was known that expression of *lin-4* forced major portions of the cuticle to remain larval rather than transforming to adult cuticle (Chalfie et al. 1981). It did so by somehow down-regulating the action of *lin-14*, a gene that produced something (soon shown to be the LIN-14 protein) maintaining cuticle cells in a larval state (Ambros and Horvitz 1984). After the first larval stage, continued production of LIN-14 protein would prevent the development of various adult cells and thus

the made the affected worms (or some of their parts) freeze in the fourth or earlier larval stages. Such worms did not achieve the adult phenotype and were sterile (Ambros and Horvitz 1984, 1987). Although the network of interactions controlling switches required for the transition from larval stage to stage and to the adult was already known to be fairly extensive and complex (Ambros 1989), it was clear that expression of *lin-4* was required from the first larval stage forward to downregulate *lin-14* and achieve adulthood. Furthermore, it was also known at least by 1987 that even when *lin-4* was transcribed and had its effects, the *mRNA* of *lin-14* remained abundant in the cell.

As of 1989, the relevant genes had not yet been molecularly characterized and there were no suitable methods yet for isolating their DNA (Ruvkun et al. 2004). The Ruvkun lab spent considerable effort exploring appropriate methods to fill in the gaps. That year, the Ruvkun lab managed to find the approximate location of the *lin-14* gene and show that *lin-4* probably downregulated production of the LIN-14 protein by blocking translation of its mRNA rather than preventing formation of that RNA (Ruvkun and Giusto 1989). They also had begun to characterize the regulatory relationships among *lin-4*, *lin-14* and several other genes (see, e.g., Arasu et al. 1991, Wightman et al. 1991). By 1993, they had accomplished a major step in deciphering the mechanism involved. The Ambros lab had localized *lin-4* to an intron of another gene and had shown that *lin-4* produces a small untranslated RNA, but does not encode a protein (Lee et al. 1993). [Talk through the phone call of June 11, 1992?] The Ruvkun lab then showed that there were seven short stretches of 21-22 nucleotides of the 3' UTR of the lin-14 mRNA that matched, albeit imperfectly, with the lin-4 RNA (Wightman et al. 1993). Note in particular, that the claims about base-pairing were caculational, rather than experimental: the complementary base pairing of the 22-nucleotide *lin-4* RNA, if it occurred, should yield characteristic loop structure. Note, too, that the region of the mRNA from *lin-14* in which the base pairing should occur is the regulatory region posterior to the coding

material and that the effect of the putative base pairing was to block, somehow, the translation of that mRNA either before or after it reached a ribosome. They also showed that no other part of the lin-14 mRNA was needed for lin-4 activity to downregulate production of the LIN-14 protein (Wightman et al. 1993).

This unexpected regulatory mechanism was quite startling at the time. But for a variety of reasons, it did not seem to be general. Some skepticism remained because of important experimental difficulties. In fact, the publication of the 1993 paper was delayed because the reviewers wished to see the results of a decisive confirmatory experiment testing the hypothesis that the putative duplexes of lin-4 and lin-14 RNAs actually occurred and had the expected regulatory effects. It turned out that the available mutations and/or the RNAs and their duplexes in the then-available experimental conditions were too unstable for that experiment to work. As it was, it took another three years to produce a confirmatory paper (Ha et al. 1996), demonstrating the necessity of the seven *lin-14* complementary sites for *lin-4* down-regulation and reporting that the desired duplexes had at least been obtained in vitro. By 1999, the Ambros lab showed that the blockage occurred after initiation of translation (Olsen and Ambros 1999); it was subsequently shown that effect was to block elongation of the LIN-14 polypeptide. For a brief review of the relevant mechanisms see (Leung and Sharp 2006).

Perhaps of greater importance, there were additional grounds for suspicion of the generality and importance of the finding: 22 nucleotides was smaller by at least a factor of four than any other regulatory RNA known at the time, and the known heterochronic genes seemed to be peculiar to the nematodes (e.g., fishing efforts in the databases available at the time found no homologues for *lin-4* or *lin-14* outside of nematodes), so this might be a regulatory system peculiar to them. Furthermore, RNA-RNA duplex interactions might be leftover prokaryotic interactions (e.g., like those involved in bacterial viral defenses), not

generally found in eukaryotes (Ruvkun et al. 2004, S95). Thus, even Ambros and Ruvkun treated the new regulatory mechanism as an intriguing, but probably parochial, mechanism in an unusual regulatory network, apparently an ad hoc mechanism of uncertain generality. Until the additional unknown links in the complex regulatory network had been pinned down or until there was some indication that the new mechanism was employed elsewhere, it was not clear whether this intriguing mechanism had major importance. And Ruvkun drew an interesting moral from the difficulty of achieving clear confirmatory experiments. The accumulation of several diverse lines of confirmatory evidence is what really did the job: "elegance in molecular genetics is aesthetically pleasing, but scientifically overrated" (Ruvkun et al. 2004, S94)! [Comment on the emblematic importance of this quotation for the (re)entry of natural history in handling the molecules under investigation. The aesthetic ideal of subsumption under pleasing physics-type laws is under challenge.]

So far, the discovery of miRNA is a nice story, but it shows little evidence of the sea change in which we are interested. Notice the timeline published by another of the pioneers. What really sprung things loose was the discovery that a previously-known lethal gene in *C. elegans* (already designated by the name *let-7*) that produced a 21-nucleotide RNA (Reinhart et al. 2000). Loss of *let-7* activity caused numerous types of cells to remain in a larval state; the miRNA produced by the *let-7* gene interacted with 39 sites on the 3' UTRs of five distinct heterochronic genes (one of which is our old friend lin-14). Like *lin-4*'s 22-nucleotide RNA base pairing with these sites (but now the evidence was much stronger, *let-7*'s 21 nt RNA played a key regulatory role in switching off translation of the mRNAs of all five of these genes, each of which helped maintain the larval condition of at least some cells (Reinhart et al. 2000). When those UTRs were spliced to a reporter gene, base pairing of the UTR with the *let-7* RNA turned off the reporter. More dramatically, let-7 proved to be ubiquitous in

animals, turning up in DNA data bases and in searches for RNA for corals, annelids, arthropods, mollusks, vertebrates, and many other organisms, but not in yeasts, E. coli, or arabidopsis. Yet further, in Drosophila and in zebrafish, as in C. elegans, let-7 function was required for temporal regulation. Here, at last, was the demonstration that the peculiar mechanism found in C. elegans was of fundamental importance. One year later, these 21-23 nucleotide RNAs received their formal name – microRNAs – with the publication of three papers in *Science* reporting on biochemical and genomic analysis of about 60 miRNAs in C. elegans plus 14 in drosophila and 19 from human HeLa cells, together with a "Perspectives" report by Gary Ruvkun explaining the background and significance of the findings (Lagos-Quintana et al. 2001, Lau et al. 2001, Lee and Ambros 2001, Ruvkun 2001).

Due to time limitations I can only take one more step before breaking off this narrative and turning to its morals. But that step is critical for drawing some of the morals that I hope you will take home with you today. It concerns how research proceeded on miRNAs (which was, I suggest, pretty much how it had to proceed). As it has turned out, the ~22 nucleotide miRNAs are normally made by very particular mechanisms – they are produced by complex processing in two stages from much longer double-stranded RNAs, which are processed in the nucleus to yield the (~70-nucleotide) pre-miRNAs that are exported from the nucleus and processed into miRNAs in the cytoplasm, where, if they are complexed into a RNA-Induced Silencing Complex, interact with other RNAs, usually mRNAs. [The original double-stranded RNAs, primary-miRNAs (pri-miRNAs) are formed using inverse repeat sections in the RNA products of miRNA genes and subjected to the first round of processing in the nucleus. In animals, this processing normally yields a ~70 nt-long product, called a pre-miRNA, shaped like a hairpin, with a typical stem-loop structure that has a some 'bumps' on the stem and with two unpaired nucleotides at each end, which are used in processing the pre-miRNA. (The story is slightly different in plants, but the

principles are similar.) In the cytoplasm, the pre-miRNA is processed into a complex of several proteins, the RNA-Induced Silencing Complex within which, under specific circumstances, it is cleaved by an enzyme of the Dicer family to yield the active product.] The miRNA then serves as a recognition device by means of which the RISC complexes with a target molecule. A great variety of things can happen to that target molecule, depending on the biochemistry of the cellular context and the specific molecular context in which the RISC bind to the RNA to which the miRNA is complementary. The simplest example is this: If the complementation of the miRNA and its target is perfect (especially if the target is in a coding region) the RISC slices the target apart, thus, if it is an mRNA, preventing initiation of transcription. If, on the other hand, the complementation is partial, depending on what is happening at other nearby sites in the target RNA, the miRNA may modulate, or contribute to the modulation of the likelihood that the target RNA will be translated. It may also modulate the likelihood of yet further processing interactions, including the likelihood of the target molecule binding to additional miRNAs. An enormous amount of biochemical analysis and genetic analysis has been done on all this processing. One result is that it allows detection of double stranded RNA molecules that are complexed for processing and another is that miRNAs can be found by chasing dicer molecules and other molecules to which miRNAs and their precursors are complexed. The inverse repeat motifs that make the double-stranded RNAs also can be pursued by computational genomics, as can many other relevant signatures. These tools are complex and have proved very important.

An enormous explosion of work on miRNAs was triggered by the demonstration of their ubiquity and their importance. What I want to stress about this explosion is how dependent it was on understanding the role played by *extrinsic* materials in figuring out what miRNA is and what it does. Exploration of the biogenesis and processing miRNAs and of their functions depended enormously on indirect experimentation, computational exploration

of sequence data bases, and exploitation of experimental tools that concerned the proteins with which they complex in addition to work on miRNAs as such. Counting the number of miRNAs involved in some interaction, recognizing their potential targets, identifying the various modulatory effects they can have (which often involve non-additive interactions between several miRNAs when they bind on a series of neighboring sites on a particular mRNA) – these all involve combining specific knowledge of the miRNAs with extrinsic information about the other molecules involved in a network which is typically composed of at least scores of molecules. miRNAs can target a virtually unlimited number of distinct sequences. They often act as rheostats, with quantitative effects that depend on the number of sites that are occupied at key moments, when the cell or the molecular region is in an appropriate state to respond to the occupation of those sites in a particular way. Thus, by proceeding from sequence information alone, it is virtually impossible to determine whether or not a particular sequence of ~22 nucleotides is a microRNA; indeed, the same sequence, produced by dicer [the molecule that slices double-stranded RNAs into ~22 nucleotide lengths and separates the two resulting strands] in different contexts from a double-stranded RNA can enter the RNA interference pathway and act quite differently than a miRNA. Thus, one must know the processing history of the sequence, or whether it complexes in the right ways with appropriate processing complexes or appropriate targets, or a staggering amount about its cellular context even to determine whether that sequence is a miRNA or a molecule that acts in the RNAi pathway. And even from a list of the potential targets of that sequence, it is virtually impossible to determine the function or functions of an miRNA from its sequence, for its targets can be scattered over the entire genome and respond differentially in different cell types, at different developmental stages, or in the presence of different exogenous chemicals. For these reasons, the intrinsic biology of ~22 nucleotide RNAs is uninformative about whether those RNAs function as miRNAs or, if they do, what

function(s) they perform. The answer to such questions depends on the enormous network of regulatory units, composed of RNAs, proteins, and numerous other cellular structures and their states or conditions. In short, it is only in the context of the network or system of interacting regulatory devices that a proper analysis of microRNAs provides a reasonably full account of how they work, how they have evolved, and what tasks they perform. This exemplifies the sea change that my collaborators have found nearly everywhere we looked.

Back to the Central Dogma: General Conclusions

As you can tell, we have begun to draw morals from the case study material. So let me stand back and use the Central Dogma as a vehicle to start pulling some threads together. We saw that Sydney Brenner's views evolved regarding how the genetic program determines whatever, exactly, it determines and that he came to recognize the necessity of separating the role of the DNA as a bearer of information from the role of that information in informing, determining, or specifying the features of the organism. He recognized, I believe, that the genes did not – and could not – contain the full information about what to do in constructing an organism. Rather, in the right (complex, dynamical, cellular) setting, the genome contains, at best, instructions about how to respond when the information it contains is unpacked in specific contexts and settings. But, I argue, the contingencies that go into when and how that information is unpacked, and how it is processed before or during its use cannot be specified by DNA alone. This makes the notion of information somewhat problematic. For today's purposes, however, it is best to let that pass and turn to another point.

The very same information yields different results in different contexts. This is as true at the cellular level as it is for the sorts of information that we humans in our daily lives. And at the cellular level, there is a second sort of heritability beyond genetic heritability – the heritability of structures and contexts. The irreversibility of differentiation in most

multicellular organisms is one marker of this. The loss of omnicompetence in development is coordinate with the acquisition of specialized competences. It is very rare for cell daughters to become less specialized than their cell mothers – which is to say that there are systematic interconnections involved at several levels that make preservation of omnicompetence rare and important. In genomics and systems biology, some understanding of the hierarchical networked interconnections underlying these facts is required before one can assess what it is that, in different contexts, can be read out of DNA instructions. Even at the level of making a protein from a given polypeptide, this turns out to be a big issue, as is shown dramatically by the phenomenon of gene sharing (Piatigorsky 2007), in which exactly the same exons, read out from a specific gene, yield exactly the same sequence of amino acids, i.e., the identical polypeptide, in different cells – but yield different proteins. In humans, for example, the same polypeptide that yields an aldehyde dehydrogenase in the liver yields a lens crystallin in the lens of the eye (and at least 20 other instances have been documented in humans in lens crystallins alone). Nearly all animals with eyes of one sort or another exhibit similar gene sharing, but with taxon-specific genes, making distinct proteins from identical polypeptides that are biochemically very diverse, coopted to make lens crystallins as well as some other product. [Since gene sharing is not widely known, I have added a slide illustrating how phylogenetically widespread it is and how diverse the molecules are that are subject to gene sharing (Piatigorsky 2003). All of these examples come from the best-studied example of gene sharing, that of lens crystallins. Other examples extend the reach of gene sharing considerably.] Gene sharing and epigenetic heritability of cell type illustrate phenomena (of which there are many more) that cannot be accounted for on standard versions of the hardened Central Dogma. This recognition drives me to a very strong reading of the consequences of the steps that Brenner took in his thinking about this topic.

As sketchy as it has been, my account of the history of work on miRNA and the suggestions I made about the sea change that took place when it became obvious that microRNAs can be found all over the map – this time, all over the map means from bacteria to fungi to plants to animals – and that they perform hugely different functions in different cells, even in the same organism, functions that can not be read off their nucleotide sequences. These findings show that miRNAs are an appropriate vehicle to use in trying to pin down in concrete detail what sorts of things are going on in the ongoing transformation of the molecular biological sciences. By using Brenner's distinction between the information in the genes or genome and what controls the instantiation of that information in a given organism, we encounter the puzzling, indeed alarming, magnitude of the gap between information and organism, between information and function. Following this path will, I believe, make the magnitude of the gap much clearer and the concrete details of what goes into making an organism – the importance of the regulatory network and its properties and of the hierarchy of pre-existing structures in the cell that are modified to make different types of cells – will make it more pressing than ever before to learn how the regulatory structure that processes genetic information and the hierarchical structure of the cell and organism (both of which also draw on exogenous information) contribute to the formation of the organism. For biologists, I suspect the concrete issues will seem more important than general and abstract reflections on those details. And yet, such abstract considerations are very powerful. In hindsight, it is not too hard to argue that many scientists were guided into error by their strong interpretations of the Central Dogma – and we need to understand how it has misled as well as how it has helped us as we try to work out appropriate methodologies to understand how such fascinating regulatory units as microRNAs work and what they do. By standing back far enough, we can see perhaps, now that we have climbed up the ladder provided by the Central Dogma, it is time to throw that ladder away.

Putting the point more generally, we have returned in these considerations to old biological questions about determination of specificity – how are traits specified?, how are different traits specified at different stages of an organisms life history?, how are species-specific traits specified? In all of these cases, it appears that the control of gene action, so called (that is, of what genes, or even genomes, as such, do) has supplied an answer that runs against the implicit ideology of Mendelian and early molecular genetics: The dimensionality of the problem of specification of the details of the organism is much greater than the information contained in the genes (at least if that information is interpreted in terms of the genetic code and how the code must be unpacked to make an organism). By themselves, the genes don't do much of anything. When the environment is contained within a certain range, genes in various specific heritable cellular contexts with stable molecular contents specify traits rather exactly, with varying degrees of built-in contingency. The tightness of specification that made it appear to geneticists that genes do the main job of specifying the organism demonstrates the enormous evolutionary accomplishment that went into stabilizing complex organismal traits and maintaining complex lineages of organisms with stable forms. But that accomplishment is as much an accomplishment of stabilizing regulatory systems and organismal structures that process the information in the genes as it is the accomplishment of stabilizing the codification of that information. And, though I haven't presented much argument in this direction today, I believe that the sort of mechanistic holism that I adumbrated at the beginning of today's talk is a plausible vehicle, perhaps the most plausible one available, for analyzing this evolutionary accomplishment.

The miRNA story is far from alone in providing a good vehicle for getting at the very large issues raised by these considerations, but it is exceptionally valuable for the light it sheds on what must be understood – and also how much is not yet understood, if we hope to provide a useful account of the ways in which genetics has consumed its older ideology and

forced us to abandon the harder-edged sorts of genetic determinism on which it was largely built. The study of miRNAs and all the other regulatory molecules that are now being found also provide a marvelous vehicle for exploring François Jacob's key insight, embodied in his famous metaphor that evolution is a tinkerer, a *bricoleur* (Jacob 1977), and the consequences of that insight as we think about sorts of experimental work required to explore how well (or poorly) integrated the systems are in which the tinkering, the *bricolage,* takes place.

Thank you for your attention.

References

Ambros V.R. and Horvitz, H.R., 1984, 'Heterochronic mutants of the Nematode Caenorhabditis elegans', Science, 226: 409-419.

Ambros V.R. and Horvitz, H.R., 1987, 'The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events', Genes and Development, 1: 398-414.

Ambros V.R., 1989, 'A Hierarchy of Regulatory Genes Controls a Larval-to-Adult Developmental Switch in C. elegans', Cell, 57: 49-57.

Ankeny R.A., 1997, 'The Conqueror Worm: An Historical and Philosophical Examination of the Use of the Nematode Caenorhabditis elegans as a Model Organism', History and Philosophy of Science. Pittsburgh: University of Pittsburgh.

Ankeny R.A., 2000, 'Fashioning descriptive models in biology: Of worms and wiring diagrams', Philosophy of Science, 67(Suppl.): S260-S272.

Ankeny R.A., 2001, 'The natural history of Caenorhabditis elegans research ', Nature Reviews Genetics, 2: 474 479.

Ankeny R.A., 2007, 'Wormy logic : model organisms as case-based reasoning'. In Creager A.N.H., , Lunbeck, E. and Wise, M.N. (eds.), Science Without Laws : Model Systems, Cases, Exemplary Narratives, Durham, NC: Duke University Press, 46-58.

Arasu P., Wightman, B.C. and Ruvkun, G.B., 1991, 'Temporal regulation of lin-14 by the antagonistic action of two other heterechronic genes, lin-4 and lin-28', Genes and Development, 5(1825-1833).

Brenner S., 1974, 'The genetics of Caenorhabditis elegans', Genetics, 77: 71-94.

Burian R.M., 2007, 'On microRNA and the need for exploratory experimentation in post-genomic molecular biology', History and Philosophy of the Life Sciences, 29(3): in press.

Burian R.M., 2009, 'Selection Does Not Operate Primarily on Genes'. In Arp R. and Ayala, F.J.J. (eds.), Contemporary Debates in the Philosophy of Biology, Oxford: Wiley-Blackwell, in press.

Chalfie M., Horvitz, H.R. and Sulston, J.E., 1981, 'Mutations that lead to reiterations in the cell lineages of C. elegans', Cell, 24: 59-69.

Crick F.H.C., 1958, 'On protein synthesis', Symposia of the Society for Experimental Biology, 12: The Biological Replication of Macromolecules: 138-163.

Crick F.H.C., 1970, 'Central dogma of molecular biology', Nature, 227: 561-563.

de Chadarevian S., 1998, 'Of worms and programmes: Caenorhabditis elegans and the study of development', Studies in the History and Philosophy of Biology and Biomedical Sciences, 29: 81-105.

de Chadarevian S., 2004, 'Mapping the worm's genome: tools, networks, patronage'. In Gaudilliére J.-P. and Rheinberger, H.-J. (eds.), From Molecular Genetics to Genomics: The Mapping Cultures of Twentieth-Century Genetics, London and New York: Routledge, 95-110.

Elliott K.C., 2007, 'Varieties of Exploratory Experimentation in Nanotoxicology', History and Philosophy of the Life Sciences, 29(3): 313-336.

Félix M.-A., 2008, 'RNA interference in nematodes and the chance that favored Sydney Brenner) ', Journal of Biology, 7(9): article 34.

Ha I., Wightman, B. and Ruvkun, G.B., 1996, 'A bulged lin-4/lin-14 RNA duplex is sufficient for Caenorhabditis elegans lin-14 temporal gradient formation', Genes and Development, 10: 3041-3050.

Jacob F., 1977, 'Evolution and tinkering', Science, 196: 1161-1166.

Lagos-Quintana M., Rauhut, R., Lendeckel, W. and Tuschl, T., 2001, 'Identification of novel genes coding for small expressed RNAs', Science, 294: 853-858.
Lai E.C., 2003, 'microRNAs: Runts of the Genome Assert Themselves', Current Biology, 13: R925–R936.
Lau N.C., Lim, L.P., Weinstein, E.G. and Bartel, D.P., 2001, 'An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans', Science, 294: 858-862.
Lee R.C., Feinbaum, R.L. and Ambros, V.R., 1993, 'The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14', Cell, 75: 843-854.
Lee R.C. and Ambros, V.R., 2001, 'An extensive class of small RNAs in Caenorhabditis elegans', Science, 294: 862-864.
Leung A.K.L. and Sharp, D.H., 2006, 'Function and localization of microRNAs in mammalian cells', Cold Spring Harbor Symposia on Quantitative Biology, 71: 29-38.
Maupas E., 1900, 'Modes et formes de reproduction des nematodes', Archives de Zoologie expérimentale et générale, 3e sér., 8: 463-624.
Morange M., 2006, 'The Protein Side of the Central Dogma: Permanence and Change', History and Philosophy of the Life Sciences, 28: 513-524.
Morange M., 2008a, 'The Death of Molecular Biology?' History and Philosophy of the Life Sciences, 30(1): 31-42.
Morange M., 2008b, 'What history tells us XIII. Fifty years of the Central Dogma', Journal of Biosciences, 33(2): 171-175.
O'Malley M.A., 2007, 'Exploratory experimentation and scientific practice: Metagenomics and the proteorhodopsin case', History and Philosophy of the Life Sciences, 29(3): 337-360.
Olby R.C., 1975, 'The Protein Version of the Central Dogma', Genetics, 79(Suppl): 3-14.
Olsen P. and Ambros, V.R., 1999, 'The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after initiation of translation', Developmental Biology, 216: 671-680.
Piatigorsky J., 2003, 'Gene Sharing, Lens Crystallins and Speculations on an Eye/Ear Evolutionary Relationship', Integrative and Comparative Biology, 43: 492-499.
Piatigorsky J., 2007, Gene Sharing and Evolution: The Diversity of Protein Functions, Cambridge, MA: Harvard University Press.
Reinhart B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R. and Ruvkun, G.B., 2000, 'The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans', Nature, 403: 901-906.
Ruvkun G.B. and Giusto, J., 1989, 'The Caenorhabditis elegans heterochronic gene lin-14 encodes a nuclear protein that forms a temporal switch during development', Nature, 338: 313-319.
Ruvkun G.B., 2001, 'Glimpses of a tiny RNA world', Science, 294: 797-799.
Ruvkun G.B., Wightman, B.C. and Ha, I., 2004, 'The 20 Years It Took to Recognize the Importance of Tiny RNAs', Cell, S116(2004): S93-S96.
Sapienza C., 2009, 'Selection Does Operate Primarily on Genes: In Defense of the Gene as the Unit of Selection'. In Arp R. and Ayala, F.J.J. (eds.), Contemporary Debates in Philosophy of Biology, Oxford: Wiley-Blackwell, in press.
Strasser B.J., 2006, 'A World in One Dimension: Linus Pauling, Francis Crick and the Central Dogma of Molecular Biology', History and Philosophy of the Life Sciences, 28: 491-512.
Sulston J.E. and Brenner, S., 1974, 'The DNA of Caenorhabditis elegans ', Genetics, 77(95-104).

Wightman B.C., Bürglin, T.R., Gatto, J., Arasu, P. and Ruvkun, G.B., 1991, 'Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development', Genes and Development, 5: 1813-1824.

Wightman B.C., Ha, I. and Ruvkun, G.B., 1993, 'Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans', Cell, 75: 855-562.