Identity, Reduction, and Conserved Mechanisms: Perspectives from Circadian Rhythm Research

William Bechtel
Department of Philosophy, Center for Chronobiology, and Interdisciplinary
Programs in Cognitive Science and Science Studies
University of California, San Diego

After briefly flourishing as a characterization of the relation between mind and brain in the 1950s (Place, 1956; Smart, 1959), the type identity theory was eclipsed for the rest of the century, supplanted by functionalism. Perhaps the most influential argument against identity theory and for functionalism was the claim that mental processes are multiply realized: the same mental phenomenon, for example hunger or pain, is realized in radically different ways in different brains, such as those of octopi and humans (Putnam, 1967). Although philosophers sometimes alluded to biological examples of multiple realizations, it was for the most part taken to be an obvious truth not requiring empirical support. As a clincher, it was sometimes noted that human brains differ (for example, in their number of neurons or the details of the wiring diagrams between neurons) and yet people share many beliefs (e.g., belief in the truth of the multiple-realization claim). Since one thing cannot be identical to two or more realizations, the alleged identity between psychological phenomena and brain processes seemed clearly refuted. An even broader conclusion was reached based on this rejection of the identity theory: psychology is autonomous from and should be pursued independently of neuroscience. This prescription fit well with the Zeitgeist in cognitive psychology and the emerging interdisciplinary field of cognitive science in the 1970s, when there were few tools available to relate findings in neuroscience to cognitive phenomena (Bechtel, Abrahamsen, & Graham, 1998). Cognitive research typically employed psychological evidence (e.g., reaction times) and computational modeling (often involving manipulation of symbol structures), but not evidence directly about the brain processes involved.

Starting in the 1980s, however, psychologists and neuroscientists began to build bridges and increasingly linked their inquiries together (LeDoux & Hirst, 1986). The advent of techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) played important roles in developing the field of cognitive neuroscience. Something very much like identity claims began to appear in cognitive neuroscience research as types of psychological processes were linked to specific types of brain processes. More recently something even more disturbing for those convinced by multiple realization arguments has begun to develop. Whereas PET and fMRI were the focus of much attention because they could be employed in the study of human brains, many neuroscience techniques, such as recording from implanted electrodes and inducing lesions, are far more invasive and not applied to humans for ethical reasons. But neuroscientists are increasingly relating findings on other species using more invasive techniques to those procured with PET and fMRI on humans, generating the very sort of type identity claims the advocates of multiple realization claimed were impossible.[1]

---

[1] In earlier work, McCauley and I (Bechtel & McCauley, 1999; McCauley & Bechtel, 2001) characterized such identity claims as *heuristic* so as to emphasize that they are typically advanced as assumptions that then guide

How do cognitive neuroscientists cope with the claimed multiple realization of mental processes? For the most part, they simply ignore the philosophers' objections and pursue their science. So the more relevant question is how should philosophers, especially those who might feel the tug from the arguments for multiple realization, respond to these developments in cognitive neuroscience? To address this issue, I will first place it in a broader framework. The issue is not unique to psychology and neuroscience, but arises in the relation of physiology to chemistry, or indeed wherever two disciplines, one of which is viewed as more fundamental, offer different perspectives on the same phenomenon.

Both philosophers and scientists employ the term *reduction* in characterizing relations between the results of higher-level and more basic-level inquiries that are supposedly jeopardized by multiple realization, but they typically understand reduction quite differently. In the first section, I will describe an understanding of reduction provided by the framework of mechanistic explanation that fits with the pursuits scientists label reductionistic. While this will help resolve some confusions that often intrude into the philosophical literature (e.g., that identity claims are themselves reductive), it will not resolve the problem of multiple realization. Rather it allows for its reformulation in terms of multiple types of mechanism generating the same type of phenomenon. It is indeed the case that there are differences between the mechanisms in different species that result in what are treated as the same phenomena. In the second section I will take up this issue directly, first arguing that when the same standards of typing are applied to phenomena as to realizations, in most instances what were viewed as the same phenomenon are in reality very similar but nonetheless different phenomena. Second, I will consider what happens when one uses a coarser grain to type neural phenomena. Then types might range across species and enable scientists to claim that the same type of mechanism in different species produces the same type of phenomenon. As has long been noted in philosophy, the notion of similarity must be appropriately constrained before it can be useful. I will introduce the notion of *conserved mechanism* that is widely employed in biological research to constrain appeals to similarity between mechanisms, and in the final section, discuss why this is such a powerful concept in contemporary biology.

Throughout this discussion I will appeal to research on circadian rhythms as an exemplar as this is a field in which the issues concerning multiple realization, conservation of mechanism, and identity, can be clearly illuminated. A further virtue of this exemplar is that scientists now have a rather detailed understanding of the mechanisms responsible for circadian rhythms, which is not the case for cognitive phenomena. Circadian rhythms are the endogenous oscillations of approximately 24-hours exhibited in a vast range of biological organisms that regulate many physiological phenomena (e.g., basic metabolism) as well as behavioral phenomena (e.g., locomotion). These endogenously generated rhythms typically vary somewhat from 24 hours, and so must be entrained by cues such as light in order to enable the regulated functions to remain appropriately linked to time of day on our planet. These rhythms also exhibit another distinctive feature for biochemical processes—they are largely unaffected by temperature, which

---

subsequent investigation. This perspective answers one classic objection to the traditional identity theory—that the evidence for identity never goes beyond demonstrating correlation. We also showed that their value as heuristic claims depends on the coarseness of the grain they use to group different realizations into the same type, thereby maintaining a type-type identity relation.

typically has a major effect on the speed of chemical reactions. The basic reactions responsible for these rhythms are performed within cells, although in multi-cellular organisms such as mammals synchronization between individual cells plays an important role in generating regular behavior (Welsh, Takahashi, & Kay, 2010). As I will show, research over the past few decades has provided substantial understanding of the mechanisms responsible for these rhythms in all orders of living organisms.

Identity, Reduction, and Mechanistic Explanation

It is important to clarify at the outset what it is to claim identity and what role identity claims play in scientific research. An identity claim appears to assert a relation, but it is a distinctive relation in that it is between something and itself. For an identity statement to be informative, what are linked in a statement of identity must be two descriptions of the same entity or a process. When an fMRI study relates a cognitive operation with a brain region (e.g., recognizing faces with activity in the fusiform face area, Kanwisher, McDermott, & Chun, 1997), when an ERP study identifies a cognitive operation with a change in electrical potentials (e.g., semantic processing with an N400, a negative deflection around 400 milliseconds post-stimulus, Kutas & Federmeier, 2000), or a single-cell recording study identifies detection of a feature of a stimulus with increased action potentials of a given populations of neurons (e.g., motion detection with action potentials in MT, Britten, Shadlen, Newsome, & Movshon, 1992), the intended claim is that the neural activity is the performance of the cognitive operation: increased activity in the fusiform gyrus is the recognition of faces, etc. (Such claims might be faulted as incorrect, perhaps because the cognitive operation involves activity in a broader range of neural processes. But that criticism simply posits a different identity claim, and does not challenge that practice of making identity claims.)

Since the same thing is not at two different levels, identity claims do not cross levels. For example, they do not relate something and its parts. The brain activity in question is not at a lower level of organization than the cognitive activity. Rather, the very same thing is being picked out in two different ways by the two descriptions. The two descriptions may pick out the process from different perspectives—as an operation that is part of a larger network of cognitive operations (a more functional characterization) or as the operation performed in a system made of particular parts (a more structural characterization). But it is still the operation that is being identified with itself. Hence, identity claims are not themselves reductionistic.

Why, then, is a neuroscience description of a process so often treated in philosophical commentary as at a lower level than a psychological characterization? In part this may be simply a vestige of the philosophical treatment of levels in accounts of reduction which analyzes them in terms of an assumed hierarchy of sciences, with physics at a lower level than chemistry, chemistry at a lower level than biology, etc. (Oppenheim & Putnam, 1958). This conception is rendered problematic by the fact that physics itself deals with phenomena ranging from constituents of atoms to astronomical. It is strained to claim that cosmological phenomena are at a lower level than chemical reactions or physiological processes (for further discussion of levels, see Craver, 2007). But the treatment of identity claims as reductive may also reflect the promise of what might be achieved by some identity claims: whereas psychological explanation typically bottoms-out with operations in an information processing mechanism, the identification of these

operations with operations within brain systems opens up the prospect of explaining how these operations are performed (e.g., by modifying the electrical potentials in a group of neurons so as to alter the patterns of action potentials they generate). On this construal, an identity claim is preparatory to a reduction, but it is not itself reductive.

To flesh out the relation between identity and reduction, it is necessary to examine what reduction involves. Philosophers have traditionally construed reduction as a relation between theories, with theories treated as linguistic structures and reduction the deduction of the reduced theory from the reducing theory (Nagel, 1961). One of the few worked out cases that seems to fit this perspective is the derivation of the phenomenological gas laws from laws of statistical mechanics. A number of virtues for such reductions have been proposed, from unification of the sciences to sharing evidence between theories or using one theory as a heuristic guide for developing other theories (Wimsatt, 1976; for discussion, see Bechtel & Hamilton, 2007). Within biology, philosophers have considered whether Mendelian genetics, characterized in terms of Mendel's laws, can be derived from molecular genetics, with many arguing that reduction, characterized in this way, fails (Hull, 1974). The reference to Mendel's laws suggests that laws are plentiful in biology and the challenge for reduction is how to relate laws in a deductive relation. However, most explanations in biology do not appeal to laws (this is in fact part of the problem of deriving Mendel's laws from molecular biology, as molecular biology is not presented in terms of laws).

Instead of explaining phenomena by showing how they fall under laws, biologists more commonly speak of explaining a phenomenon by discovering and characterizing the mechanism responsible for it. Beginning in the 1990s, a number of philosophers set out to analyze what biologists mean by mechanism and how they go about adducing mechanisms to explain particular phenomena. Although there are differences between the various accounts that have been advanced (Bechtel & Richardson, 1993/2010; Glennan, 1996, 2002; Machamer, Darden, & Craver, 2000; Bechtel & Abrahamsen, 2005; Darden, 2006; Craver, 2007), a key aspect of all of them is the claim that a mechanism consists of parts that perform operations which are organized so as to generate the phenomenon of interest. Identifying parts and operations requires decomposing a mechanism—taking it apart (actually, or in analysis) either structurally into its parts or functionally into its operations. One important part of this task is localizing operations in the parts that perform them. Localization claims are identity claims between the parts characterized in terms of their physical constitutions (e.g., proteins characterized in terms of their sequence of amino acids and patterns of folding) and parts characterized in terms of the operations they are involved in (e.g., catalyzing particular chemical reactions). These identity claims are one level down from the initial identity claim linking a phenomenon to the mechanism producing it.

For the characterization of the parts and operations of a mechanism to explain a phenomenon, one must be able to understand how those parts and operations work together to produce that phenomenon. A key aspect of this is to understand the operations, which are processes of change (for which reason Machamer, Darden, and Craver refer to them as *activities*). A further aspect is to understand how the parts are organized and their operations coordinated. In accounts such as Machamer, Darden, and Craver's, the functioning mechanism is assumed to involve sequential execution of operations "from start or set-up to finish or termination conditions." Many

biological mechanisms, however, involve non-sequential organization, where the individual reactions are non-linear when described mathematically and operate in an environment that is open to flows of energy. Under these conditions mechanisms can exhibit complex dynamic behavior. To understand this dynamic behavior, Abrahamsen and I advocate integrating tools of mathematical modeling (including tools for analyzing complex systems) and mechanistic research, generating what we call *dynamic mechanistic explanations* (Bechtel & Abrahamsen, 2010). One characteristic of dynamic mechanistic explanations is the use of mathematical models to understand how complex behaviors such as endogenous oscillations are generated by the mechanism.

The pursuit of mechanistic explanation is well illustrated in research on circadian rhythms. The first step was to locate the mechanism. In research on mammals the strategy that proved productive was to trace the projections whereby light could entrain rhythms. This guided Moore (1973) to locate the mechanism in a small nucleus (consisting of 8,000 to 10,000 neurons on each side of the brain) in the hypothalamus that is situated above the optic chiasm and so known as the *suprachiasmatic nucleus* (SCN). Subsequent research showed that lesions to this structure disrupted rhythms, and at least some circadian rhythms could be restored by transplanting a functioning SCN into a ventricle (Silver, LeSauter, Tresco, & Lehman, 1996).

The identification of the SCN revealed where circadian rhythms were generated, but not how they were generated. That required discovering the mechanism generating oscillations within individual cells. A pioneering step was the discovery of a gene in *Drosophila* in which mutations resulted in either shortened or lengthened periods or arrhythmic behavior. Konopka and Benzer (1971) named the responsible gene *period* (*per*). Once researchers could clone *per* and measure concentrations at different times of day, Hardin, Hall, & Rosbash (1990) proposed the mechanism portrayed in Figure 1 in which *per* is transcribed into its mRNA, which is transported to the cytoplasm where it is translated into the protein PER. The protein eventually is transported back into the nucleus and by a process, at the time unspecified, inhibits its own transcription until most of it is degraded. Then transcription and translation begin again. The result is that concentrations of PER oscillate over the course of a day. One important question was whether a mechanism of this sort would produce sustained oscillations, or eventually settle into an equilibrium condition. Goldbetter (1995) developed a mathematical model based on the reactions in the mechanism Hardin et al. proposed that showed that such a mechanism, with biologically realistic parameters, would indeed produce sustained oscillations. The combination of such modeling with the mechanistic research that identified the components of the mechanism exemplifies dynamic mechanistic explanation. Since 1995, homologs of *per* have been found in mammals (e.g., mice) and many additional clock genes have been identified in *Drosophila*, mammals, and indeed in plants, fungi, and bacteria, and the operations they perform in the clock articulated. Further modeling efforts have both identified the parameters under which continued oscillations could occur and examined consequences of various ways the mechanism might be disrupted or altered (see Bechtel & Abrahamsen, 2010, for discussion).
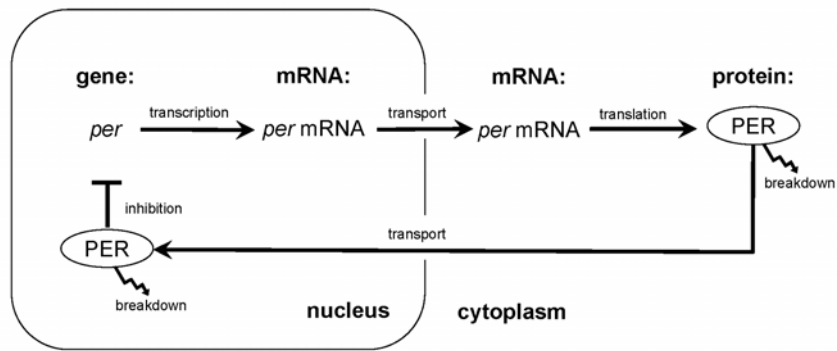
Figure 1. Hardin et al.'s (1990) proposed feedback mechanism for generating circadian oscillations in PER concentrations

What is most significant for current purposes is that mechanistic explanations are inherently reductive—decomposing a mechanism into its parts and operations involves going to a lower level of organization. Genes such as *per* are parts of the intracellular mechanism that oscillates and thus at a lower level than the cells (or the SCN). There is typically further organization of sub-parts within the parts of the mechanism (e.g., the order of the nucleotides within *per*), but that is not the focus in explaining circadian oscillations. Rather, the explanation consists in describing the various operations and the organization imposed on the parts (the feedback loop illustrated in Figure 1). It is this organization that renders the mechanism an entity over and beyond the collection of parts and operations which, in isolation, do not oscillate. Moreover, oscillatory behavior is not just an additive combination of the operations of the parts, but results from the non-sequential organization of operations that when characterized mathematically are non-linear. In this respect, circadian oscillation is an emergent phenomenon, but emergence and reduction, thus construed within the context of mechanistic explanation, are not exclusive alternatives but perspectives that must be integrated in an explanation.

The Multiple Realization Objection viewed from a Mechanistic Perspective

Having articulated the mechanistic perspective and how it applies to research on circadian rhythms, I return to the multiple-realization objection. It might appear to resurface with a vengeance in the mechanistic framework—if different mechanisms could generate the same phenomenon, that phenomenon is multiply realized. In our commercial society, examples of multiple realizations seem to be extremely widespread—we characterize an appliance in terms of the function it performs, and we know that different manufacturers can produce appliances out of different components, performing different operations, organized in different ways. Recalling that multiple realization was viewed as a problem for identity claims required for reductionistic research, there would seem to be a direct way of presenting these problems within the framework of mechanistic explanation. Indeed, it would seem to constitute a death knell to mechanistic explanation itself. The identity claims linking a phenomenon with a particular mechanism or an operation within the mechanism with a part of it would seem to be jeopardized. Why are biologists so keen on mechanistic explanation when it appears so clearly problematic philosophically?

Even if the claim multiple realization were non-problematic, its defenders might acknowledge one reason to understand how a given phenomenon is produced by a particular mechanism. When a manufactured system needs to be repaired, the repair must be appropriate to a specific realization. If different manufacturers use different parts and operations, organized in different ways, and a part of one mechanism breaks, it is critical to choose as a replacement a part that meets the specification of the broken part, not a part that might work in one of the other mechanisms that would produce the same phenomenon. Likewise, when a biological mechanism is diseased, the treatment must be appropriate to the specific mechanism operative in that organism. Acknowledging this role for mechanistic understanding, those enamored with the multiple-realization argument would focus their objections on the claim that accounts of a mechanism *explain* the phenomenon, contending that at best they explain how the phenomenon was produced in one particular case. And, assuming alternative realizations arise even between members of a species, the proponent of multiple realization would seem to be entitled to deny that an accounts of how a mechanism generates the phenomenon in one organism is explanatory of the phenomenon within the species.

However, the case for multiple realizations is not nearly as straightforward as this suggests. First, the context of engineered products may be rather misleading about the ease of developing multiple mechanisms that can produce the same phenomenon. In the design context, there is often a premium on developing alternative mechanisms that can produce the same phenomenon as well as factors to maintain commonalities between different realizations. Social processes help insure basic commonalities among different manufacturer's products (e.g., road design keeping cars within size limits and fuel supply systems limiting the choice of fuels) even as patent protections require that competing manufacturer's develop their own designs. Even so, multiple realizations don't realize all the same functionality. This is demonstrated by the fact that consumers often turn to specifications in deciding which item to purchase, expecting them to differ. Any dishwasher may perform basic dish washing, but some models may clean more thoroughly, consume less energy, operate more quickly, or create more noise. This reveals an important consideration too often ignored in making the case for multiple realization. In claiming that hunger or pain are equally realized in octopi and humans, Putnam adopted a very loose criterion of sameness of function. There are huge differences in the behavior between hungry octopi and hungry humans, differences that are in part a reflection of the different constitution and organization of the responsible mechanisms in different species. The same point applies to the differences between members of a species—individual differences can generally be detected (even if often glossed over in behavioral research that tends to focus on means and use variability only to assess statistical reliability).

The differences between realizations is well-illustrated in research on circadian rhythms. As I noted above, circadian clocks are found in all orders of life. In all species in which the responsible mechanisms have been studied (except cyanobacteria), the core of the mechanism involves a transcriptional feedback loop that, in addition to the negative loop shown in Figure 1, employs a positive loop, as shown for *Drosophila* in Figure 2. This second loop involves the transcription and translation of CLOCK (CLK), which forms a dimer with a second protein, CYCLE (CYC), which, despite its name, is produced constitutively. The CLK:CYC dimer binds not only to the promoter for *per* and its dimerization partner *tim*, but also to the promoter of *pdp1ε*, whose protein binds to *clk*'s promoter creating the positive feedback loop. Comparable

pairs of feedback loops have been discovered in other animals, including mammals, as well as in fungi and plants. But many of the components differ. In mice homologues of many of the same genes and proteins are involved, but there are two versions of *per*, *Per1* and *Per2*, and *tim* is replaced by *Cry1* and *Cry2*.[2] In the fungus *Neurospora*, *frequency* (*frq*) functions much like *per* in *Drosophila*: its transcription is promoted by WHITE COLLAR 1 and 2 (WC1 and WC2) and once the protein FRQ is formed, two molecules form a homodimer and enter the nucleus, where they interact with WC1 and WC2 so as to inhibit their own transcription. In the plant *Arabidopsis*, *cca1* and *lhy* function like *per* and *tim*: their transcription is promoted by a homodimer of TOC1 and once they are formed, they dimerize and enter the nucleus, where they interact with TOC1 so as to inhibit their own transcription. Despite these and other differences, all the organisms exhibit endogenous circadian rhythms that are entrainable by light and temperature compensated. The circadian clock seems to be multiply realized.
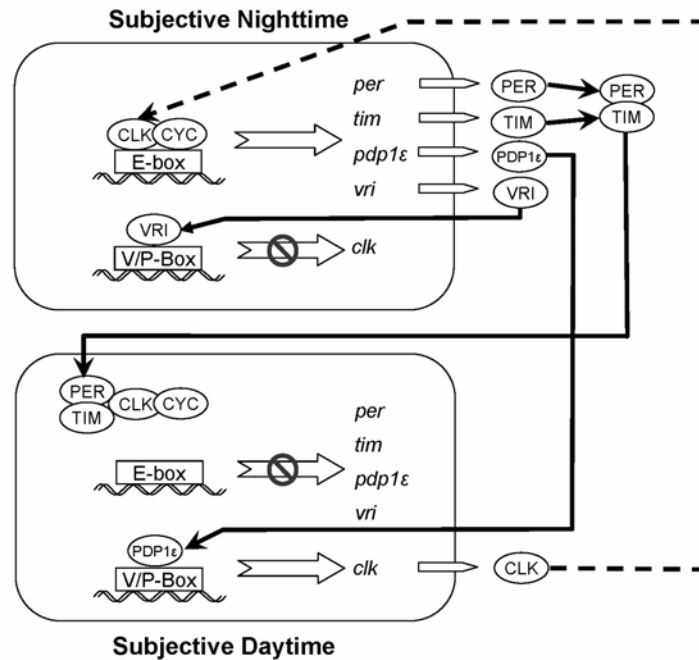

Figure 2. A more complete account of the Drosophila oscillator. The large open arrows indicate whether the promoter turns gene expression on or off. The smaller open arrows represent the combined processes of gene transcription in the nucleus, transport to the cytoplasm, and translation in the cytoplasm.

There are, however, important behavioral differences between the clocks of different species that result from these differences in parts. First, in *Drosophila* the PER:TIM dimer is at its highest concentration during the night, whereas in mammals the PER:CRY dimers are maximal during the day. To advance the *Drosophila* clock, light during what the fly experiences as late night expedites the degradation of TIM, thereby releasing the inhibitory effect on the dimer on the CLK:CYC dimer earlier than usual. But if a similar mechanism were used in mammals (perhaps affecting CRY rather than TIM), it would continually degrade the dimer during the day when it should be increasing in concentration. Accordingly, a different entrainment process is employed, one by which light promotes the transcription and translation of the mammalian *Per* genes. This entails that different molecules are required to sense and communicate the presence of light (in

---

[2] Whereas in insects gene names are usually printed in lowercase italics, in mammals the first letter is capitalized. Protein names are commonly printed in uppercase letters.

Drosophila a version of CRY and in mammals, melanopsin). Second, the plant clock shows a fundamental difference from those in animals. Plant clocks continue to operate in constant light and under constant illumination theirs oscillations are modulated by the wavelength and intensity of light (Somers, Devlin, & Kay, 1998). Finally, the differences in constitution of the different clock mechanisms requires differences in how they are linked to the various physiological and behavioral processes and thereby how they regulate these processes.

So far I have addressed the multiple realization challenge by considering how differences in the mechanism manifest themselves in behavioral differences, thereby questioning whether it is the same phenomenon that is in fact realized. In developing the question of grain, Mundale and I (Bechtel & Mundale, 1999) primarily took the opposite tactic by considering how neuroscience often follows the lead of psychology in treating the realizing mechanisms as the same despite apparent differences. Focusing on how brain areas are typically delineated by neuroscientists, we considered how areas are identified as the same in different species despite substantial changes in the structures themselves. Thus, MT is construed as the locus of motion detection in monkeys and humans despite substantial differences in its size and its location. The identity claim is maintained (MT is the locus of motion detection) by counting somewhat different brain structures as forming one type that is identified as responsible for the somewhat variable instantiations of the same behavior type.[3] This underlies the widespread use of some species in neuroscience and other fields of biology as model systems for others. I turn to this practice of using a mechanism in one species as a model for another in the next section.

Model Systems, Conserved Mechanisms, and Multiple Realization

The reliance on model systems is extremely widespread in biology. *Drosophila* have served not only as model systems for the study of circadian rhythms, but also for the study of transmission genetics in the early 20th century and, beginning in this century, for the study of sleep. As geneticists sought a more molecular understanding of genes and gene expression, the fungus *Neurospora* and the bacteriophages that infect the bacterium *Escherichia coli* played fundamental roles. Typically these species are not studied because of intrinsic interest in them, but because they are convenient to study (due to their short lifespan, ease of biochemical manipulability, lack of complications presented by gene duplications found in target organisms etc.), and are sufficiently similar to the target species (often *homo sapiens*) to be informative. The appeal to similarity between model and target is extremely important, as that is the basis for extrapolation. In many contexts, the extrapolation is only the beginning of the research process: using the model as a guide, research on the target mechanism may substantiate some of the similarities, but also reveal differences (Bechtel, 2009). As differences are revealed, the model often continues to provide a guide in the quest to identify other ways of accomplishing the same or related functions (for example, *cry* was first determined to figure in entrainment of circadian rhythms to light in *Drosophila*, but once the mammalian variants of *Cry* were found not to figure

---

[3] Philosophers might resist this move, allowing that it might be epistemically useful to scientists to proceed in this manner, but that the realizations are nonetheless ontologically different and so provide multiple realizations. One answer to this is to follow the move above and note that, in such cases, so are the phenomena realized. But a further move is to maintain an ontological focus and consider what makes for a *type* of realizer. The question then turns on whether philosophical intuitions (these particular realizers cannot be of the same type) or scientific practices are better guides to the relevant ontological types. Absent a compelling argument otherwise, it seems reasonable to follow the scientist and try to understand how they identify what they take to be relevant types in nature.

in entrainment in mammals, research was directed to identifying another candidate, a process that culminated in the discovery of melanopsin).

The question is how the similarity required to count two mechanisms or their parts as of the same type is assessed. It cannot be assessed directly by comparing the parts, operations, and organization of the two mechanisms because researchers are often investigating the mechanism in the model system in lieu of studying the target system. Even when research is later conducted on the target, it is not directed at establishing or discounting the utility of the model. The similarity is anticipated or inferred, not demonstrated. What biologists often appeal to is evolutionary descent from a common ancestor such that the mechanism in question is *conserved*. A mechanism does not have to be frozen and not altered to count as conserved; evolution involves, in Darwin's phrase, *descent with modification*. But because modifications are made from a common origin, researchers expect significant similarities, and these similarities are the basis for treating the mechanism in the different organisms as falling in the same category or type. (Conservation involves more than descent, since ultimately all organisms on earth appear to have descended from a common origin, a prokaryotic cell. Additional tools are invoked for identifying types, such as type of molecule or type of mechanism. One approach is to identify common sequences of amino acids forming proteins, counting proteins as potentially homologous when they exhibit significant sequence similarity. Below I will consider another way of typing mechanisms in terms of shared functional organization. The important factor that descent introduces is the expectation that the types identified are likely to be robust so that additional research is likely to unveil additional commonalities amongst the type.)

The mechanism for circadian rhythms within animals illustrates both aspects of descent with modification invoked in the previous paragraph: there are important commonalities but also noteworthy differences. To understand the role of both, I will elaborate on some of the historical process of discovery of the circadian mechanisms in different species. The *Clock* gene was first discovered in mice by generating and screening for circadian mutations; this search culminated in identifying a mutation that produced long circadian periods in mice when heterozygous and loss of rhythms when homozygous (Vitaterna et al., 1994). Three years later the same laboratory cloned *clock* and predicted it possessed a DNA-binding region that would enable it to perform the intermediate operation revealed in *Drosophila* research between the accumulation of the PER protein and the inhibition of *per* transcription (King et al., 1997). Subsequently Gekakis et al. (1998) showed that mammalian CLOCK, together with its dimerization partner BMAL1, could bind to the promoter box on the *Drosophila per* gene. Since the CLOCK used in this study was procured from mammals but had effects in *Drosophila*, researchers inferred that *Drosophila* contained a homolog to mammalian *Clock*, an hypothesis confirmed by Darlington et al. (1998). If the mechanisms was conserved, though, there should be mammalian homologues of *Drosophila per* and *tim* and already in 1997 two groups (Sun et al., 1997; Tei et al., 1997) found a first mammalian homologue of *per*. In the same year one of the groups found a second homologue (Albrecht, Sun, Eichele, & Lee, 1997) and the following year yet another was identified (Zylka, Shearman, Weaver, & Reppert, 1998).

The last example begins also to point to differences that appear in descent lineages. One of the most common ways differences originate is through the duplication of genes with some duplicates continuing to perform the same task (*mPer1* and *mPer2*) while others undergo change

and end up being recruited for different tasks (*mPer3*). But some differences are more major. Recall that in *Drosophila* TIM is the dimerization partner of PER. In the same period when mammalian homologues to *per* were discovered researchers also found a mammalian homolog to *tim* and initially suggested that it functioned just as *tim* did in *Drosophila* (Sangoram et al., 1998). But further research revealed a major difference: in constant darkness (or constant light) neither the *mTim* mRNA nor the mTIM protein oscillated, and in response to a light stimulus, mTIM levels increased whereas dTIM levels decreased. Even more significant differences were soon discovered. In *Drosophila* entrainment to light is mediated by a cryptochrome (*cry*) that is conserved from the blue-light photoreceptors in plants. The search for mammalian homologues of *cry* yielded two each in mice and humans. Their proteins, however, oscillated in a circadian manner even in constant darkness, whereas the homolog in *Drosophila* only oscillated when given a light stimulus, appropriate for the role it had already been assigned in the entrainment process (Todo et al., 1996). Moreover, in mammals kept in darkness, loss of CRY1 resulted in shortened circadian periods, loss of CRY2 resulted in elongated periods, and loss of both CRY1 and CRY2 eliminated circadian rhythms (van der Horst et al., 1999). These results together suggested a role for CRY1 and CRY2 in the central clock mechanism, not just in entrainment, and Griffin, Staknis, and Weitz (1999) produced evidence that CRY1 and CRY2 were dimerization partners of PER1 and PER2. They were thus recognized as playing the role TIM played in *Drosophila*.

Once again, these differences in the mechanism might seem to support multiple realization. In considering this case briefly above, I raised one line of response to this—that these differences generated behavioral differences so that it was not the same phenomenon being generated by the difference mechanisms. But the biologist's view of these cases is different—despite the differences, the mechanisms and the phenomenon are viewed as belonging to the same type insofar as they are conserved. The modifications produced in phylogeny are modifications in mechanisms comprising the same type.

While the claim of conservation of the circadian mechanism within animals is widely accepted, the standard view is that the circadian clocks in the different orders of life are not conserved but are of independent origins and have converged on a common design. But there is reason to accept conservation in this case too, which I will develop after presenting the reasons many opt for convergence of different clock mechanisms in the next paragraph. I briefly introduced one reason descent is viewed as important—when two mechanisms are descended from a common ancestor, one expects that the commonalities are deep, not superficial. A finding about one is likely to generalize to the other, either by being true of both, or indicating compensating similarities and differences (if two forms of an ancestor are formed and diverge, they will still show deeper commonalities, and if a functional component is lost, it will be replaced by something else). What is at stake in this debate over convergence versus conservation is whether this assumption is justified.

The case for the independent origin of clocks begins with the fact that, as discussed above, clocks in animals, fungi and plants are constituted by distinct components. The fact that they are nonetheless organized in the same way (as seen in Figure 3) and all serve to coordinate physiological and behavioral activities with time of day in the local environment is taken as evidence for evolutionary convergence. An additional step in the argument for independent

origins is provided by research on the cyanobacterium *Synechococcus elongates*. Initially it too seemed to depend upon a transcriptional feedback loop, but when Nakajima et al. (2005) incubated the three proteins implicated in the cyanobacterial clock, KaiA, KaiB, and KaiC, with ATP, oscillations with a period of 22 hours were sustained over several days without genes to be transcribed and translated. The oscillations involved the successive phosphorylation and dephosphorylation of KaiC at two (perhaps three) sites. KaiC manifests both autokinase and autophosphatase activity. Binding of KaiA to KaiC stimulates autophosphorylation of KaiC first of a tyrosine (T432) and then of an adjacent serine (S431). The second phosphorylation permits binding of the KaiA and KaiC complex with KaiB, which blocks further phosphorylation of KaiC. The autokinase operation then dominates, with dephosphorylation occurring first at T432 and subsequently at S431 (McClung, 2008). From findings such as these, many scientists have concluded that the circadian clock has evolved separately in the different orders of life and that the organizational similarity between them is due to convergent evolution (Young & Kay, 2001; Rosbash, 2009). If descent from a common source is the basis for assessing conservation and hence identity of mechanism, there would seem to be independent realizations of circadian timekeeping.
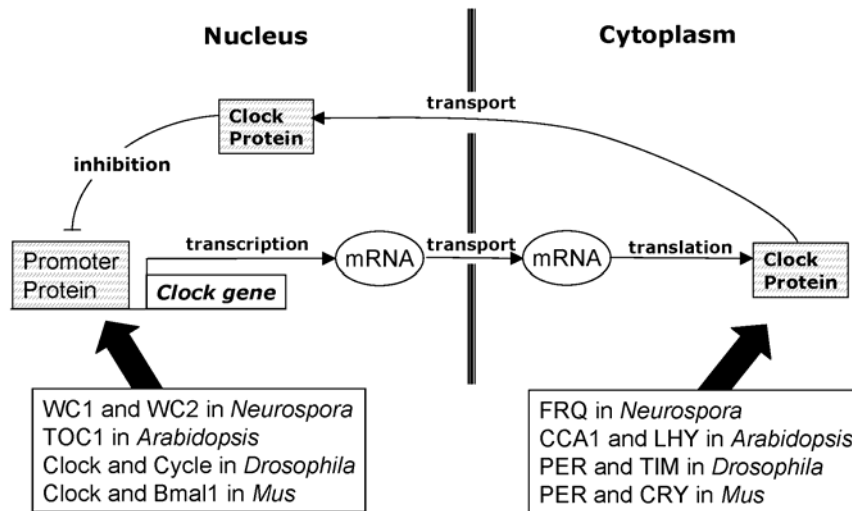

Figure 3. Schematic representation showing similar operations and organization in the clock mechanisms in different organisms, with different proteins performing the operations.

Despite the prominent differences between clock constituents in the different orders of life, there are some intriguing similarities between them that encourage rethinking the claimed independent evolution of the clocks. First, the promoters in animals (PER), fungi (WC1 and WC2) and plants (LHY and CCA1) share a region known as a PAS domain that facilitates dimerization. They also exhibit similarities in the region around the PAS domain (Young & Kay, 2001). Second, the *cry* genes found in *Drosophila* and mammals are homologues of genes found in plants, and in *Arabidopsis* there are two *Cry* genes that figure in blue-light signaling to the clock. Moreover, there is evidence that cryptochromes are present in cyanobacteria (Cashmore, 2003; Tauber, Last, Olive, & Kyriacou, 2004). This suggests that the entrainment function of *cry* might be primary, and its mammalian function a modification. Finally, variants of casein kinase 1 figure in the critical step of phosphorylating clock proteins in fungi, plants, and animals (although they

also figure in many other cell signaling processes). The constitution of the different clocks shares more than is commonly indicated in arguments for convergent evolution.

To argue for descent with modification, though, requires not just pointing to some constitutive similarities but explaining how so many components could be changed in clocks if they are part of a descent tree. Addressing this requires that we focus on what is the most striking similarities between clocks in different orders of life—that except for cyanobacteria, they all involve positive and negative transcriptional feedback loops. The convergent evolution account proposes that different evolutionary lineages all arrived at the same solution to the challenge of maintaining an approximately 24-hour oscillation. But a different possibility that is at least as plausible is that this solution was arrived at once, and it remained stable, while different genes and proteins came to perform the various operations within the organizational framework. On this scenario, what is conserved is the organization, giving rise to what Ralph Greenspan (personal communication) terms *functional homology*. On this scenario one might expect both wholesale substitutions but also some remnants of ancestral forms, such as the PAS domains, the CRY proteins and the kinases identified in the previous paragraph. The cyanobacterial clock would initially seem to be the major exception to this proposal since its organization is not the same. Greenspan, however, suggests the possibility that transcription feedback loops might have first been appended to the phosphorylation loop in the cyanobacterial clock and retained as they provided for increased robustness. Indeed, there is evidence that such transcription feedback loops are present in cyanobacteria and while not essential to it generating circadian oscillations, may provide stability (Mackey & Golden, 2007). Some, such as Tauber (2004), propose tracing the cyanobacteria clock to even earlier evolutionary stages, pointing to evidence that during the evolution of early life the earth rotated much more quickly (as fast as once every four hours) and proposing how metabolic and cell oscillations (which have periods in the four hour range) evolved to protect organisms from damaging effects of UV radiation. He proposes that these then provided a template for the later evolution of circadian clocks.

The question of whether the circadian clocks in the different orders of life are conserved variants of a common ancestor or the product of convergent evolution between different lineages is unresolved. What is important for my purposes is why the question matters. If the different clocks are products of convergent evolution, then researchers would expect there to be limits to the extrapolations that can be made from one to another as the similarities will be limited to what is required to fulfill the same evolutionary function (e.g., keeping circadian time). Likewise, if two variants of a manufactured product (e.g., an operating system) were truly independently created and only had to meet the same (or similar) desiderata, we would not expect further research to reveal deeper similarities—only those required to satisfy the desiderate. On the other hand, if there were a common designer, we would expect various vestiges of the ancestry to show up (and when they do so, that would provide evidence at the two designs were not arrived at independently). When similarities between mechanisms are observed in biology, researchers often hope that further research will yield even more understanding. Thus, they investigate whether they are homologous, and if they are, treat the mechanisms as of the same type and use research on one to guide research on the other.

Conclusions

Identifying phenomenon with a mechanism or an operation within a mechanism with one of its parts are identity claims. These identity claims play critical roles in the development of mechanistic explanations in a variety of disciplines, including those devoted to understanding cognition. Scientists continue to make such identity claims despite philosophical objections based on the contention that phenomena (or operations within mechanisms) are multiply realized and hence not identifiable with a given mechanism (or a given part of a mechanism). Employing a strategy Mundale and I introduced, I have argued that the appearance of multiple realization stems from employing different grains in characterizing phenomena and the mechanisms responsible for them (or operations and the parts that perform them). Distinguishable mechanisms only generate the same phenomenon when a much coarser grain is employed for the phenomenon than the mechanisms. By employing as restrictive a grain for the phenomena as for the mechanisms in cases where the mechanisms differ, the phenomena can be distinguished as well.

In most contexts, researchers employ a coarser grain for typing both phenomena and mechanisms as they are interested in identifying the type of mechanism responsible for a given type of phenomenon so as to use information about one member of the type to guide understanding of other members. In the example I used throughout this paper, circadian rhythms are characterized as approximately 24 hour oscillations that are entrainable by light or other cues and are temperature compensated. On first appearances, the responsible mechanisms in different animals, or across the orders of life, seem to be different as they differ in the components and, in the case of cyanobacteria in their organization. If this were the case, there would be relatively severe limits in what researchers could extract from one organism to another. When common descent is found, however, researchers expect deeper similarities so that new discoveries made on one organism will provide guidance for research on the other. Conservation provides a basis for treating the mechanisms as belonging to a common type.

I illustrated this use of evolutionary descent to identify types of biological mechanisms in the case of circadian rhythms. Within animals, there is wide agreement that the clock mechanism in insects and mammals, the two exemplars most widely investigated, is conserved. The fact that mammals have multiple versions of *Per* does not undercut treating the mechanisms as of a common type. In fact, it brings forward a reason for focusing on *Drosophila*: without the multiple forms of the gene, one can more readily study the operations performed by the gene. Even the substitution of *Cry* for *tim* does not undercut the use of *Drosophila* as a model for mammals, since biologists are accustomed to considering both commonalities and variations in conserved mechanisms. The comparison across the orders of life poses greater challenges, since there are fewer homologous parts in the clocks. But the fact that they share a common organization may indicate functional homology (same organization due to common descent) and if so provides a basis for treating all circadian clocks as conserved and hence members of the same type. Then one could speak of a common type of clock mechanism and use the clock mechanism in a given order as a model for understanding those in other orders. Underlying this is an identity claim between a type of clock mechanism and the type of circadian phenomenon.

I have focused on circadian rhythms in large part because the research into the responsible mechanisms is quite far advanced and the practices of the biologists well established. But the lessons regarding identity extend further, both to biology and to the cognitive sciences. In developing mechanistic explanations, scientists are interested in types that are relatively coarsely characterized, in part because they are interested in including in the types model systems that can be useful for elucidating the type. Thus, in neuroscience, investigators interested in the human brain commonly characterize types so as to include primates, frequently other mammals, and sometimes other phyla such as birds that have a forebrain. Increasingly they are including invertebrates such as insects since these are providing useful models for studying various psychological phenomena such as learning, memory, and sleep. What is generally taken as crucial is characterizing types and treating them the same is commonality of descent. Differences within the type are expected. Type identity claims are core to the practice of mechanistic explanation in biology and are not jeopardized by the philosophical claims of multiple realization.

References

Albrecht, U., Sun, Z. S., Eichele, G., & Lee, C. C. (1997). A differential response of two putative mammalian circadian regulators, *mper1* and *mper2*, to light. *Cell, 91* (7), 1055-1064.

Bechtel, W. (2009). Generalization and discovery by assuming conserved mechanisms: Cross species research on circadian oscillators. *Philosophy of Science, 76*, 762-773.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421-441.

Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*.

Bechtel, W., Abrahamsen, A., & Graham, G. (1998). The life of cognitive science. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 1-104). Oxford: Basil Blackwell.

Bechtel, W., & Hamilton, A. (2007). Reduction, integration, and the unity of science: Natural, behavioral, and social sciences and the humanities. In T. Kuipers (Ed.), *Philosophy of science: Focal issues* (pp. 377-430). New York: Elsevier.

Bechtel, W., & McCauley, R. N. (1999). Heuristic identity theory (or back to the future):   The mind-body problem against the background of research strategies in cognitive neuroscience. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st Annual Meeting of the Cognitive Science Society* (pp. 67-72). Mahwah, NJ: Lawrence Erlbaum Associates.

Bechtel, W., & Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science, 66*, 175-207.

Bechtel, W., & Richardson, R. C. (1993/2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge, MA: MIT Press. 1993 edition published by Princeton University Press.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *The Journal of Neuroscience, 12*, 4745-4765.

Cashmore, A. R. (2003). Cryptochromes: Enabling plants and animals to determine circadian time. *Cell, 114* (5), 537-543.

Craver, C. F. (2007). *Explaining the brain: What a science of the mind-brain could be*. New York: Oxford University Press.

Darden, L. (2006). *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution*. Cambridge: Cambridge University Press.

Darlington, T. K., Wager-Smith, K., Ceriani, M. F., Staknis, D., Gekakis, N., Steeves, T. D., et al. (1998). Closing the circadian loop: CLOCK-induced transcription of its own inhibitors *per* and *tim*. *Science, 280* (5369), 1599-1603.

Gekakis, N., Staknis, D., Nguyen, H. B., Davis, F. C., Wilsbacher, L. D., King, D. P., et al. (1998). Role of the CLOCK protein in the mammalian circadian mechanism. *Science, 280* (5369), 1564-1569.

Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis, 44*, 50-71.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*, S342-S353.

Goldbeter, A. (1995). A model for circadian oscillations in the *Drosophila* Period protein (PER). *Proceedings of the Royal Society of London. B: Biological Sciences, 261* (1362), 319-324.

Griffin, E. A., Jr., Staknis, D., & Weitz, C. J. (1999). Light-Independent Role of CRY1 and CRY2 in the Mammalian Circadian Clock. *Science, 286* (5440), 768-771.

Hardin, P. E., Hall, J. C., & Rosbash, M. (1990). Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature, 343* (6258), 536-540.

Hull, D. L. (1974). *The philosophy of biological science*. Englewood Cliffs, NJ: Prentice-Hall.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J. Neurosci., 17* (11), 4302-4311.

King, D. P., Zhao, Y., Sangoram, A. M., Wilsbacher, L. D., Tanaka, M., Antoch, M. P., et al. (1997). Positional cloning of the mouse circadian *Clock g*ene. *Cell, 89* (4), 641-653.

Konopka, R. J., & Benzer, S. (1971). Clock mutants of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences (USA), 89*, 2112-2116.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4* (12), 463-470.

LeDoux, J. E., & Hirst, W. (Eds.). (1986). *Mind and brain: Dialogues in cognitive neuroscience*. Cambridge: Cambridge University Press.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1-25.

Mackey, S. R., & Golden, S. S. (2007). Winding up the cyanobacterial circadian clock. *Trends in Microbiology, 15* (9), 381-388.

McCauley, R. N., & Bechtel, W. (2001). Explanatory pluralism and the heuristic identity theory. *Theory and Psychology, 11*, 736-760.

McClung, C. R. (2008). Comes a time. *Current Opinion in Plant Biology, 11* (5), 514-520.

Moore, R. Y. (1973). Retinohypothalamic projection in mammals: A comparative study. *Brain Research, 49*, 403-409.

Nagel, E. (1961). *The structure of science*. New York: Harcourt, Brace.

Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., et al. (2005). Reconstitution of Circadian Oscillation of Cyanobacterial KaiC Phosphorylation in Vitro. *Science, 308* (5720), 414-415.

Oppenheim, P., & Putnam, H. (1958). The unity of science as a working hypothesis. In H. Feigl & G. Maxwell (Eds.), *Concepts, theories, and the mind-body problem* (pp. 3-36). Minneapolis: University of Minnesota Press.

Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology, 47*, 44-50.

Putnam, H. (1967). Psychological predicates. In W. H. Capitan & D. D. Merrill (Eds.), *Art, Mind and Religion* (pp. 37-48). Pittsburgh: University of Pittsburgh Press.

Rosbash, M. (2009). The implications of multiple circadian clock origins. *PLoS Biology, 7* (3), e1000062.

Sangoram, A. M., Saez, L., Antoch, M. P., Gekakis, N., Staknis, D., Whiteley, A., et al. (1998). Mammalian Circadian Autoregulatory Loop: A *Timeless* Ortholog and *mPer1* Interact and Negatively Regulate CLOCK-BMAL1-Induced Transcription. *Neuron, 21* (5), 1101-1113.

Silver, R., LeSauter, J., Tresco, P. A., & Lehman, M. N. (1996). A diffusible coupling signal from the transplanted suprachiasmatic nucleus controlling circadian locomotor rhythms. *Nature, 382* (6594), 810-813.

Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review, 68*, 141-156.

Somers, D. E., Devlin, P. F., & Kay, S. A. (1998). Phytochromes and cryptochromes in the entrainment of the *Arabidopsis* circadian clock. *Science, 282* (5393), 1488-1490.

Sun, Z. S., Albrecht, U., Zhuchenko, O., Bailey, J., Eichele, G., & Lee, C. C. (1997). RIGUI, a putative mammalian ortholog of the Drosophila *period* gene. *Cell, 90* (6), 1003-1011.

Tauber, E., Last, K. S., Olive, P. J. W., & Kyriacou, C. P. (2004). Clock gene evolution and functional divergence. *Journal of Biological Rhythms, 19* (5), 445-458.

Tei, H., Okamura, H., Shigeyoshi, Y., Fukuhara, C., Ozawa, R., Hirose, M., et al. (1997). Circadian oscillation of a mammalian homologue of the *Drosophila period* gene. *Nature, 389* (6650), 512-516.

Todo, T., Ryo, H., Yamamoto, K., Toh, H., Inui, T., Ayaki, H., et al. (1996). Similarity Among the *Drosophila* (6-4)Photolyase, a Human Photolyase Homolog, and the DNA Photolyase-Blue-Light Photoreceptor Family. *Science, 272* (5258), 109-112.

van der Horst, G. T. J., Muijtjens, M., Kobayashi, K., Takano, R., Kanno, S.-i., Takao, M., et al. (1999). Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature, 398* (6728), 627-630.

Vitaterna, M. H., King, D. P., Chang, A.-M., Kornhauser, J. M., Lowrey, P. L., McDonald, J. D., et al. (1994). Mutagenesis and mapping of a mouse gene, *Clock*, essential for circadian behavior. *Science, 264* (5159), 719-725.

Welsh, D. K., Takahashi, J. S., & Kay, S. A. (2010). Suprachiasmatic nucleus: Cell autonomy and network properties. *Annual Review of Physiology, 72* (1).

Wimsatt, W. C. (1976). Reductive explanation: A functional account. In R. S. Cohen, C. A. Hooker, A. C. Michalos & J. van Evra (Eds.), *PSA-1974* (pp. 671-710). Dordrecht: Reidel.

Young, M. W., & Kay, S. A. (2001). Time zones: a comparative genetics of circadian clocks. *Nature Reviews Genetics, 2* (9), 702-715.

Zylka, M. J., Shearman, L. P., Weaver, D. R., & Reppert, S. M. (1998). Three *period* homologs in mammals: Differential light responses in the suprachiasmatic circadian clock and oscillating transcripts outside of brain. *Neuron, 20* (6), 1103-1110.