Title
Light and feeding time inputs into mammalian timekeeping mechanisms

Permalink
https://escholarship.org/uc/item/0np5t57t

Author
Gill, Shubhroz

Publication Date
2012

Peer reviewed|Thesis/dissertation
Light and Feeding Time Inputs Into Mammalian Timekeeping Mechanisms

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biology by Shubhroz Gill

Committee in charge:

Professor Satchidananda Panda, Chair
Professor Susan Golden
Professor Alexander Hoffmann
Professor Yishi Jin
Professor Reuben Shaw

2012
The dissertation of Shubhroz Gill is approved, and it is acceptable in quality and form for publication of microfilm and electronically:

_______________________________________________
_______________________________________________
_______________________________________________
_______________________________________________
_______________________________________________
_______________________________________________

Chair

University of California, San Diego

2012
# Table of Contents

Signature Page.................................................................................................................. iii  
Table of Contents ........................................................................................................... iv 
List of Figures .................................................................................................................. v  
Acknowledgements ......................................................................................................... vii 
Vita ................................................................................................................................. viii 
Abstract of the Dissertation .......................................................................................... ix 
Chapter 1. The Nature and Importance of Circadian Timekeeping in Organisms .............. 1 
Chapter 2. Rhythmic and Light-Modulated Gene Expression in the Mouse SCN .......... 10 
Chapter 3. An Infrastructural Setup for Studying Feeding Time in People ..................... 28 
Chapter 4. Insights from an Evidence Driven Approach to Studying Human Behavior .... 47 
Chapter 5. Future Directions ......................................................................................... 69 
References ..................................................................................................................... 76
List of Figures

Figure 2.1. Circadian transcription in the mouse SCN .................................. 22
Figure 2.2. Light modulated transcription in the mouse SCN ...................... 23
Figure 2.3. SCN specific loss of Lhx1 disrupts circadian rhythm of locomotor
activity ........................................................................................................... 26
Figure 3.1. Interface design for the user-facing smartphone application ...... 42
Figure 3.2. Nature of the data received for each food or drink instance ...... 43
Figure 3.3. Study workflow for subjects and researchers ......................... 44
Figure 3.4. Push notification system ............................................................. 45
Figure 3.5. Push notification system: user-facing screens ......................... 46
Figure 4.1. Age and BMI distribution of the study population ..................... 53
Figure 4.2. Self-described ethnicity of the study population ....................... 54
Figure 4.3. Initial and final BMI of the study population ............................. 55
Figure 4.4. BMI change versus initial BMI ................................................. 56
Figure 4.5. Example of pictures received from one person in one day ......... 57
Figure 4.6. Distribution of the time of eating or drinking for different individuals
.......................................................................................................................... 58
Figure 4.7. 95% interval of the time of eating or drinking for different individuals
.......................................................................................................................... 59
Figure 4.8. Intermeal time for different individuals..................................... 60
Figure 4.9. Median intermeal times for individuals..................................... 61
Figure 4.10. Feedogram................................................................................ 62
Figure 4.11. Annotations reveal patterns of food and beverage consumption ................................................................. 63

Figure 4.12. Estimating the calories and volume of each item consumed ................................................................. 64

Figure 4.13. Temporal spread of caloric intake ................................................................. 65

Figure 4.14. Temporal spread of volumetric intake ................................................................. 66

Figure 4.15. Estimation of specificity and sensitivity of the methodology ................................................................. 67

Figure 4.16. Relative activity levels for 20 individuals ................................................................. 68

Figure 4.17. Temporal spread of average activity levels ................................................................. 69
Acknowledgements

First and foremost, I would like to thank my advisor Dr. Satchidananda Panda for excellent mentoring, and for letting me pursue challenging and creative science. I would like to thank current and former members of the Panda lab, especially Dr. Luciano DiTacchio, Dr. Megumi Hirota and Hiep Le.

The work presented in this dissertation would not be possible without excellent support from my friend and wife Dr. Joyita Dutta. I would also like to thank my family members (my sister Shikhadeep Gill, my parents and my parents-in-law) for their perseverance and support.

Last but not the least, I would like to acknowledge the contributions of Sami Michishita and Alyssa Nelson, who were instrumental in carrying out the human research study.

Chapter 2 is being submitted for publication under the following title: Megumi Hatori, Shubhroz Gill and Satchidananda Panda, “Transcriptional landscape of the SCN and the role of Lhx1 as a determinant of SCN function”. The dissertation author was the co-first researcher and author of this paper, along with Dr. Megumi Hatori.

Chapter 3 will be submitted for publication and the dissertation author was the primary researcher and author of this paper.
Vita

2012  Ph.D. in Biology, Division of Biological Sciences, University of California, San Diego

2006  B. Tech. (Honours) in Biotechnology and Biochemical Engineering, Department of Biotechnology, Indian Institute of Technology Kharagpur

Publications


ABSTRACT OF THE DISSERTATION

Light and Feeding Time Inputs Into Mammalian Timekeeping Mechanisms

by

Shubhroz Gill

Doctor of Philosophy in Biology
University of California, San Diego, 2012
Professor Satchidananda Panda

Terrestrial organisms have evolved mechanisms to anticipate the environmental changes associated with diurnal variations in the environment. The primary such mechanism is the circadian timer which is a molecular clock present in every cell of mammals. In order to adjust to external variations, the circadian clock is plastic and responds to environmental stimuli. Two such stimuli are light and food timing. In this dissertation I describe two separate studies, one in mice and another in humans where we have attempted to characterize these two stimuli.
In chapter 2, I describe a series of experiments performed in mice in which we have defined the extent of circadian and light driven transcription in the master clock, namely the suprachiasmatic nucleus, residing in the hypothalamic region of the brain.

In chapter 3, I describe the development of a behavioral monitoring and data analysis system for understanding the extent of temporal spread of caloric intake in people. To our knowledge this is the first time that smartphones have been employed to characterize human behavior. Our data reveals previously unknown aspects of eating behaviors.

In chapter 4, the results obtained from the human behavioral study are discussed. Lastly, in chapter 5, I mention our strategy for future studies using the approach described in chapter 3 and 4.

Overall, this dissertation expands our knowledge of light and feeding time inputs to the mammalian timekeeping mechanisms.
Chapter 1. The Nature and Importance of Circadian Timekeeping in Organisms

Darwinian evolution provides mechanisms for generation and selection of heritable mutations that improve the survival and reproduction of organisms. Terrestrial organisms have evolved to survive in the presence of a 24-hour cycle in environmental changes such as nutrient availability, ambient temperature and light availability. These endogenous variations in biological function or characteristics that have a time period of approximately 24h and whose purpose is to allow an organism to inherently keep time independent of environmental input are known as circadian rhythms. Circadian rhythms are derived from the Latin words circa, meaning "approximately," and diem, meaning “day” and are a Darwinian adaptation for survival and fitness on a rotating planet. An organism that is able to anticipate the upcoming changes associated with these can better prepare for it and thus would have a survival advantage. Thus, it is not surprising that organisms ranging from bacteria to mammals have evolved mechanisms that help them keep time, with the time period of roughly twenty-four hours.

The key defining features of circadian rhythms are:

Endogenous nature: As distinguished from temporal variation imposed by external cues such as an imposed light/dark cycle, circadian variations persist without any external timing cue. For example a mouse, fruitfly or a plant kept in
the dark and at a constant temperature will display variation in metabolite or transcript abundance.

Time period $\approx 24h$. This is a feature that is designed around the time period of the earth’s rotation around the sun, and the chief environmental challenge that circadian rhythms evolved to overcome.

Entrainability: The endogenously generated rhythms in organisms must synchronize to external variation. In addition to evolving mechanisms to survive and reproduce optimally in an environment that changes drastically every 12h, terrestrial organisms must also continue to perform optimally in spite of the seasonal variation in the length of the duration of daylight itself. Thus, the circadian clockwork mechanism must be able to . A well-known manifestation of this property is the jetlag that humans experience when they travel across time zones and must synchronize to a different light/dark cycle, and the disruption in sleep wake cycle that occurs in the interim while this synchronization is not complete.

Temperature compensation: This is considered a defining characteristic of circadian rhythms and at the molecular level implies that the biochemical reactions must have a $Q_{10} \approx 1$ around the physiological temperature. That is, the variation in defining characteristics of the circadian rhythm such as the time period, amplitude and phase must remain unchanged if the temperature is varied
within a reasonable range around the physiological temperature of the organism. Temperature compensation is a surprising feature of circadian clocks because biochemical parameters such as rate constants of reactions and the abundance of molecules varies a great deal with temperature.

Any biological parameter that shows circadian variation, can be defined by its amplitude (one half of the difference between the maximum and minimum), the time-varying phase and the time period. The resetting of circadian clocks refers to a change in phase. Typically, resetting occurs when an organism moves from one light/dark cycle to another such as a person going from the west coast of the US to the east coast or in the periphery in response to food cues.

By studying the genetic requirements for circadian rhythms in different model organisms spanning eukaryotes and prokaryotes, it has been realized that circadian rhythms may be an example of convergent evolution. Homologs of circadian components in one organism often have no function in generating circadian rhythms in another, with the exception of the newly described peroxiredoxins described above. Remarkably, in the absence of functional equivalence of homologs, the actual structure of the system is very similar. The core clock mechanism is a transcriptional-translation feedback loop (TTFL) in which one set of proteins drives the production of its own repressors. The repressors then act on activators to suppress their function, after a short delay.
This produces rhythmic variation in the abundance of both sets of proteins, which then drives a consequent temporal variation in their downstream targets.

The ability to align systemic processes with nutrient availability and day length improves fitness and fecundity in many species. In cyanobacteria, synchrony between the photoperiod and periodicity of the endogenous rhythms improves fitness (Ouyang et al., 1998). In plants, the circadian clock regulates the transition from the vegetative to the reproductive phase and thereby determines the flowering season. As in plants, the circadian clock in insect populations native to different latitudes shows features that likely improve adaptation toward local climates (Pittendrigh et al., 1991). However, the regulatory complexity of organ systems in animals has precluded studies of whether and how circadian clocks impact broad physiological outcomes.

Recently, an intriguing discovery was made regarding peroxiredoxins (Edgar et al., 2012). It has been found that the peroxiredoxin family of proteins shows circadian variation in many different species, and thus may be a conserved molecular component of the circadian machinery. Till date, this is probably the only molecular component of the circadian clock whose homologs across various species seem to have some functional significance for circadian rhythms. Further, this has suggested that circadian rhythms may have evolved at a time when oxygen became available on earth as a response to its rhythmic diurnal availability.
In mammals, the cell autonomous molecular mechanism that forms the gears of the circadian clock is implemented in the following manner. A heterodimeric complex of Clock and Bmal1/Arntl proteins drive transcription by binding to genomic sites called E-boxes. Among the transcriptional targets of the Clock/Bmal1 complex are their own repressors: the Period and Cryptochrome families of proteins. The Period protein (Per1, Per2 or Per3) heterodimerizes with a Cryptochrome (Cry1 or Cry2) and translocates to the nucleus to repress the activity of Clock and Bmal1 complexes, with a short time delay. This is known as the negative limb of the circadian oscillator. As a result the abundance of Per/Cry and Clock/Bmal1 complexes varies sinusoidally over the course of a day. In addition, the core clock mechanism also has a so-called positive limb that is implemented as follows. The Rev-erb-α gene is a target of the Clock/Bmal1 complex that drives its circadian fluctuation over the course of a day. The Rev-erb-α (transcriptional repressor) and ROR-α (transcriptional activator) proteins then sequentially occupy genomic sites called ROR elements in the Bmal1 promoter and accordingly lead to the Bmal1 transcript being produced at high or low levels.

The circadian variation in these core clock proteins drives fluctuations in transcription of other genes that mediate specific aspects of cellular, tissue-level or organismal physiology. Some transcripts only show circadian variation in one cell type but not another, and this is an aspect discussed in a later chapter in this dissertation. For example, the CYP7A1 gene whose protein catalyzes the rate-
limiting step in the conversion of cholesterol to bile acids has circadian abundance (Noshiro et al., 2007). By controlling the amount of transcript and hence the protein, the circadian clock helps separate different aspects of physiology to different temporal niches.

Comparative time-course gene expression profiling at a high temporal resolution over the entire circadian cycle of mice that possess or lack a functioning circadian clock has demonstrated that approximately 5-10% of the transcripts in the genome show robust circadian fluctuations (Panda et al., 2002). While the identity of the transcripts varies from one tissue to another, there is some overlap between them.

The cell autonomous circadian oscillators are organized into hierarchies of tissues. The master synchronizing signals are provided by the hypothalamic suprachiasmatic nuclei (SCN). The SCN consists of about 20,000 neurons that show circadian variation in gene expression and electrophysiological activity. The mechanisms of how these cells are coupled and synchronized are discussed later. In short, the SCN provides the major resetting signal in response to changes in environmental light conditions. The peripheral clocks are the next lower level in this synchronizing hierarchy and reset their clocks in response to SCN signals (Albrecht et al., 2012). In addition, the periphery can also respond to food timing as a cue. The circadian clock tunes its properties precisely to the environmental cycles by integrating many such inputs. These inputs, which can
be food, light, temperature etc. are referred to as zeitgebers. In mammals, the principal input to the SCN is light as sensed by a special class of ganglionic photoreceptors that express the photopigment melanopsin. A neural pathway called the retinohypothalamic tract then conveys this information to the SCN. Peripheral organs such as liver, on the other hand, are additionally highly sensitive to time-of-feeding as a temporal cue.

Given how fundamentally intertwined metabolism and behaviors are to diurnal and circadian variation, it is not surprising that genetic and environmental alteration of normal circadian function causes multiple pathological consequences in people. This has been established through multiple lines of evidence.

The first line of evidence is genomewide association studies in people that have linked clock loci with disease. In one of the earliest genetic study of a human cohort with a sleep syndrome, Toh and colleagues identified a mutation in Period2 gene that underlies FASPS in these individuals (Toh et al., 2001). Period2 is a critical and core component of the circadian clock in mammals. In 2005, Xu et al discovered that a mutation in the clock enzyme CKIδ that underlies a heritable disorder called Familial advanced sleep phase syndrome (FASPS) where individuals wake up and sleep much earlier than the average person, a condition that is colloquially referred to as “larkism” (Xu et al., 2005). This demonstrates how the circadian component of sleep phases the sleep and
wake times to an appropriate time of the day. In a Finnish study, SNPs in ARNTL2 were associated with risk for clinically diagnosis of social phobia (Sipila et al., 2012). SNPs adjacent to the melatonin receptor MTNR1B and cryptochrome CRY2 have been observed to co-segregate with levels of fasting glucose and other related phenotypic characteristics in healthy individuals. The implication of the circadian clock in human disease has also led to a renewed interest in pharmacology of clock components. It was recently showed (Hirota et al., 2012; Cho et al., 2012 and Solt et al., 2012) that two clock components: Cry and Rev-erb-alpha are druggable.

The second line of evidence is the increased incidence of disease or suboptimal physiology when environmental alterations perturb normal clock function. It has been recently shown that there is a statistically significant increase in the number of acute myocardial infarctions on the days of transition from and to daylight saving time (Janszky et al., 2008). It has also been shown that combining sleep restriction with an imposed phase change in sleep/wake cycle (to model circadian disruption) leads to syndromic metabolic disease such as decreased insulin, elevated postprandial peak glucose in human volunteers in a laboratory study (Buxton et al., 2012). In a study of ~27000 individuals, it was discovered that there is an elevated risk for metabolic disease in shift workers, particularly in females (Karlsson et al., 2001). Firemen are known to have a higher lag time (for responding to medical exigencies) at certain times of the day.
versus others. The response time was least at 10am (best performance) and highest at 1am (worst performance) (Brousse et al., 2011).

The third set of data that implicate circadian clocks to normal physiology and homeostasis are studies in model organisms. Clock mutant mice develop obesity on both normal and high fat diets and exhibit hypertriglyceridemia, decreased insulin and higher glucose (Turek et al., 2005). Bmal1 knockout mice recapitulate multiple aspects of cardiovascular disease and have impaired healing responses to artery ligation (Anea et al., 2009). RNAi-mediated knockdown of Clock in the mouse VTA leads to manic-depressive symptoms (Mukherjee et al., 2009).

Taken together these data suggest that impaired clock function has deleterious consequences for normal physiology, and that rejuvenation of a dampened clock might provide therapeutic benefit under such circumstances.
Molecular circadian oscillations are prevalent in all tissues that have been examined till date. A significant fraction of the genome is under circadian control, at least as measured by transcript abundance. *In vitro* these transcriptional oscillations can even be observed in dispersed single cells from different tissue types. *In vivo*, in mammals, these cellular oscillators exist in the context of tissue and thus they must synchronize to their surrounding cells as well as respond to signals from other tissues. The major such tissue that provides synchronizing signals to other tissues are the two symmetrically located suprachiasmatic nuclei (SCN) of the hypothalamus in the brain. For this reason, the SCN is often referred to as the “master clock” located in the mammalian brain.

Melanopsin expressing retinal ganglion cells (mRGCs) in the retina of the eye detect blue light with extremely slow kinetics and then feed this information into the circadian system, thereby enabling synchrony between the organism's external environment and internal physiology. The long integration time of melanopsin photoreception that underlies the slow kinetics is a key feature that distinguishes non-image-forming mRGC mediated vision from image-forming photoreceptors such as rods and cones, because it protects circadian clocks from being reset by a short pulse of light such as a midnight lightning bolt. Melanopsin vision also has a high intensity threshold than the amount of light
available under moonlight to once again avoid resetting clocks. mRGCs project their axons to multiple regions of the brain, prominent among which is the SCN. The retinal input is integrated, by poorly understood mechanisms to shift the phase of circadian oscillations. The resetting is then conveyed to peripheral organs, again by unknown mechanisms.

The hallmark features of the hypothalamus suprachiasmatic nucleus (SCN) master oscillator are its intrinsic oscillations that produce rhythmic outputs in secreted and synaptic communication, gated response to light, and SCN specific coupling of the circadian oscillators. Although transcriptional regulation is known to underlie the core functional mechanism of the mammalian circadian system, transcriptional dynamics underlying SCN function is largely unknown.

To comprehensively identify circadianly regulated protein-coding transcripts in the SCN, we collected total RNA every 2 h over two complete days from male C57BL6/J adult mice entrained to 12h light: 12 h darkness and subsequently held under constant conditions. Using Affymetrix MOE430 high density arrays we discovered 1,667 probesets corresponding to 1,412 genes (Figure 2.1a,b) show circadian oscillation in transcript abundance (pMMCb < 0.05, pFGT < 0.05, median temporal expression > 100). Of these, genes corresponding to 122 probesets including the core clock components and their targets were previously known to be oscillating in the SCN (Ueda et al., 2002; Panda et al., 2002). Under constant darkness these rhythmic transcripts reached
the peak levels of expression at similar time of the day (within 8h) as found in previous studies (Fig. 2.1c), thus validating the method and statistical criteria used in the present study to discover rhythmic transcripts in the SCN.

Oscillations of majority of the rhythmic transcripts were specific to the SCN. Out of the 1667 rhythmic transcripts, 398 were also detected as rhythmic in the mouse liver (Figure 2.1d). This small subset of common-cyclers was enriched for known clock components and immediate output targets. Although the liver and the SCN shared many rhythmic transcripts nearly 50% of them oscillate with a peak time of expression differing by >8 h between these two tissue types. Such tissue specific nature of the circadian transcriptome suggested participation of their protein products in SCN specific oscillatory function.

The peak phase of expression of the rhythmic transcripts were distributed throughout the 24 h with 51% of rhythmic transcripts reaching the peak expression level at late subjective night (CT16-CT20) and another cluster of 22% peak at mid subjective day (CT06-CT10) coinciding with the time of the mid activity and rest of nocturnal mice (Fig. 2.1e). Such prevalence of rhythmic transcripts clustered in one major phase is characteristic of other organs and organisms. This suggested temporal regulation of RNA processing and protein turnover might support increased expression of a large number of transcripts at night. Accordingly, functional annotation of the rhythmic transcripts revealed the top two gene ontology (GO) terms for RNAs peaking during the night time related
to RNA processing and protein folding while those peaking during daytime were lysosomal or lytic vacuoles related. Upon deeper inspection, 51 probe sets related to mRNA splicing (p=4.6E-11) were synchronously expressed with a peak expression in mid subjective night. At the same time among the second highly enriched term related to protein folding (p = 2.1E-9), we detected mRNAs encoding 6 out of 8 subunits of the Tcp1 chaperone show synchronous coordinated oscillation. Such close temporal coordination of RNA processing and protein folding complexes likely support rhythmic expression and function of large number of SCN transcripts during the night time, which are subsequently recycled and degraded at daytime.

A key feature of the SCN is paracrine signaling of multiple peptidergic signals including VIP, AVP and GRP that enforces coupling of individual oscillator neurons to sustain robust oscillations over several cycles. Rhythmic expression of the peptide or their cognate receptors with peak phase of expression during subjective night likely reinforces intercellular coupling.

Light signal perceived in the retina and transmitted to the SCN via the melanopsin expressing RGCs constitutes the dominant entrainment cue to the SCN to synchronize the circadian clock with the ambient light environment. Light pulse is known to trigger extensive chromatin remodeling but only upto few dozen transcripts are known to be light modulated at night, leaving the vast majority of light induced changes and time-of-the-day effect of light in the SCN
largely unknown. We thus assessed the transcriptional changes in the SCN in response to 60 min light pulse delivered at circadian times 6, 16 and 22 representing subjective noon, late evening and early morning, when light causes no shift, phase delay or phase advance of the behavioral rhythm respectively. By comparing the levels of the transcripts in the light-exposed and dark-maintained animals at 0, 1, 2 and 4h after light pulse (a.l.p.), we identified nearly 576 transcripts whose levels changed by light by > 2 fold in response to at least one of the three light pulses. This group included more than 2 dozen known light induced genes in the SCN, many of which were also independently confirmed by qPCR, thus validating the approach and statistical cutoff criteria used. Light induced transcriptional changes are more prevalent during subjective night than in subjective daytime, thus suggesting circadian gating of response to light in the SCN that coincides with the overt circadian behavioral phase shift induced by nighttime illumination.

To test whether the light induced transcriptional changes correlate with light resetting of the overt behavioral rhythms, we measured transient changes in gene expression of a subset of transcripts by qPCR in mice with specific perturbation of retina photoreceptors. Mice lacking inner retina photopigment melanopsin show attenuated behavioral response to light, while those lacking outer retina rod/cones photoreceptors show near normal behavioral response to light. Both groups show normal circadian behavioral rhythms. Circadian gating of light responses are intact in the Opn4-/- mice, thus suggesting gating is a
property of the clock and not of the retina photopigment. Light induced transcriptional changes are severely attenuated in Opn4-/- mice while rd/rd mice lacking rod/cone photopigment show mild attenuation of light induced changes in the SCN. Light responses were completely abolished in mice lacking rod/cone and melanopsin photoreceptors (rd;opn4-/-) mice or mice with acute ablation of melanopsin RGCs (Opn4Cre/+;iDTR + DT), thus conclusively establishing light induced transcriptional changes in the mouse SCN are entirely mediated by the rod/cone/melanopsin photopigment and transmitted through the mRGCs.

Stimulus dependent resetting of the ensemble of SCN neurons involves both resetting of the individual cellular oscillators and perturbation of the cell-cell communication that modulate cellular synchrony. We reasoned that transcripts that are circadianly expressed and also light regulated might form important nodes for integrating light information into the clock. Out of 1667 rhythmic transcripts only 88 are also light modulated. In general, among the light modulated and oscillating group, transcripts that typically peak in the daytime, are induced by light at night and show no significant change in response to light during subjective day, while those peaking at night are suppressed by light at night. Hence, light pulse trigger transcriptional signature in the SCN that mimic the state of the SCN at day and might offer a molecular explanation of the light induced change in the master clock. Among the known clock components the expression of only Per genes were induced by light in a time-of-the day dependent manner. Although light induced upregulation of transcripts in the SCN
have been well described, we also found transcript levels of several genes reduced over 4 h time following light pulse. Among the light suppressed group are several genes implicated in cell-cell communication including Avpr1a which supports the long standing hypothesis that light perturbs coupling among SCN oscillators which, in turn promotes resetting the phase of the ensemble of SCN neurons.

Since there are a large number of induced and suppressed transcripts, we turned our attention to transcription factors that might be mediating these changes. Using pscan (Zambelli et al., 2009), we performed in silico binding site enrichment analysis in gene promoter regions for light modulated transcripts. 10 genes (Plxna2, Aplp2, Cdk14, Il1rap, Tmem151, Kcna1, Kif5a, Msra, Ppp2r5c and Zfp664) are genes induced by light and contain potential binding sites (score>.95) for Egr1 in their genomic 1kb upstream regions. Egr1 itself is ~4 fold induced by light. Taken together, the light-modulation of Egr1 and its potential transcriptional targets strengthens the case for Egr1 as an important transcription factor that might control light-responsive transcription in the SCN.

Although circadian oscillators are pervasive, the role of SCN as a master oscillator partly lies in its peptidergic properties. So we hypothesized SCN enriched transcription factors impart such function on the SCN. To identify SCN enriched genes we employed three step enrichment criteria to find transcripts that are enriched relative to 91 mouse tissues, 14 different neural tissues and
enriched relative to the hypothalamus transcripts in the GNF tissue atlas database (Lattin et al., 2008). This criteria identified 230 probesets representing 213 SCN enriched transcripts, many of which including Rora, Rorb, AVP, VIP, GRP, RGS16, Prokr2 were known to be SCN enriched (VanDunk et al., 2011; Prosser et al., 2007). Furthermore, a BAC transgenic reporter mouse line expressing GFP from RGS16 gene, also showed cell type specific expression in the SCN and the retina, thus validating our approach to find SCN enriched transcripts.

Lhx1 is a transcription factor enriched in the SCN, whose expression early during development of mouse development coincides with the expression of SCN enriched neuropeptides VIP and AVP. Lhx1 plays a critical role in differentiation and/or migration of a number of neuronal and renal cell types (Shawlot et al., 1995; Cirio et al., 2011) and the loss of Lhx1 in mice causes embryonic lethality (Kania et al., 2000). To test the role of Lhx1 specifically in the mouse SCN, we generated Rora-Cre;Lhx1^{loxp/loxp} double mutant mice for specific deactivation of Lhx1 in the SCN. Rora is expressed in the SCN and other brain regions, however, the overlap between Rora and Lhx1 in the brain is largely restricted to the SCN. The Rora-Cre mouse was generated by knocking in an IRES;Cre cassette 3' downstream of the Rora locus, which does not attenuate Rora function in the Rora-Cre mice (as opposed to the developmental and circadian dysfunction in Rora KO or hypomorphic staggerer mutants (Sato et al., 2004) exhibit normal development, body weight and circadian activity rhythm that is indistinguishable
from those of the WT mice. Rora-Cre expression is enriched in the SCN as RORα-Cre;Z/AP or RoraCre;LacZ\textsubscript{loxP/loxP} mice show robust expression of a Cre-dependent ALPP or LacZ expression only in the SCN region of the hypothalamus. ALPP staining of RoraCre;Z/AP mice revealed the majority of SCN neurons show only local projection with rare long range projection beyond the immediate vicinity of the SCN.

The RoraCre/Cre:Lhx1\textsubscript{loxP/loxP} (Lhx1SCN-KO) mice develop normally and show no gross defect in development or general locomotor activity. The Lhx1 mRNA level in the SCN is reduced significantly (Fig 2.3) indicating the effectiveness of Rora-Cre in SCN specific deactivation of Lhx1 gene. Their locomotor activity patterns in 12h light: 12h dark (LD) shows that they entrain normally to an imposed light:dark cycle, suggesting normal innervation of the SCN by the mRGCs.

Under constant darkness, the average locomotor activity of the RoraCre/+;Lhx1\textsubscript{LoxP/LoxP} or RoraCre/Cre; Lhx1\textsubscript{LoxP/LoxP} mice was not different from that of the WT cohorts. The RoraCre/+; Lhx1\textsubscript{LoxP/LoxP} mice show normal activity rhythm for the first 2-3 weeks under constant darkness, after which the activity consolidation deteriorates with no ~24h rhythm. The RoraCre/Cre; Lhx1\textsubscript{LoxP/LoxP} mice on the other hand exhibit a severe disruption of circadian activity rhythm under constant darkness.
The lack of circadian activity rhythm in RoraCre/Cre; Lhx1\textsuperscript{loxp/loxp} mice does not likely result from disruption of the cell autonomous circadian oscillator, as the median expression of many core clock components largely remains equivalent in the SCN of RoraCre/Cre;Lhx1\textsuperscript{loxp/loxp} and WT cohorts (Figure 2.3F). Expression of Per1 and DBP show significantly dampened rhythm with reduced peak levels and increased expression at the trough, while the median expression levels remained equivalent to WT cohorts.

We next tested the expression of SCN enriched transcripts that play a role in intercellular communication and thus help synchronize the cellular oscillators. Remarkably, the expression of VIP and Avpr1a were significantly reduced in the Lhx1 SCN KO mice. Avpr1a and the receptors for VIP are G-protein coupled receptors which use Ca2+ and cAMP second messengers for intracellular signal transduction. Additionally, expression of intracellular signaling molecules Pde7b, Creb3l1, Rasd1 and of a cell matrix associated cell-cell interaction mediator Nov are also reduced in the Lhx1 SCN KO mice. The loss of VIP (Colwell et al., 2003) or Avpr1a (Li et al., 2009; Wersinger et al., 2007) in mice causes mild period alteration and under prolonged darkness leads to loss of circadian rhythm, while Rasd1 KO mice exhibit reduced sensitivity to light induced synchronization of the SCN to the ambient LD cycle. This indicates a critical role of Lhx1 in determining the specific feature of SCN neurons that imparts coupling among neurons.
In summary, we describe the circadian, light regulated and tissue enriched protein-coding transcriptome of the mouse SCN. The results vastly expand the dynamic landscape of the SCN. A unique feature of the SCN is its tight intercellular communication imposed by a paracrine peptidergic signals largely mediated by VIP and AVP. We found expression of transcription factor Lhx1 is required for expression of a number of genes whose protein products participate in intercellular signaling. The loss of Lhx1 leads to reduced amounts or absence of its transcriptional targets that are important for intercellular communication and thereby likely attenuates the cell-cell coupling of SCN oscillators leading to loss of circadian activity consolidation.
Figure 2.1. Circadian transcription in the mouse SCN.
(a) Radar plot for probesets that are rhythmic in the SCN. Each colored circle corresponds to one probeset. The center of the colored circle from center of the figure is proportional to the logarithm of the median temporal value. The radius of a circle is proportional to the amplitude. The color and the angular location of a circle depict the phase i.e. the time of the day at which the transcript shows peak expression. (b) The peak phase of probesets for transcripts that were also present in an earlier version of the microarray platform used for SCN expression profiling were used to compare the peak phase of those transcripts in the current study with that in a previously published study. (c) The number of transcripts that are rhythmic in the liver or the SCN or both. (d) The phase difference for rhythmic probesets in SCN and liver.
Figure 2.1 (continued). Circadian transcription in the mouse SCN.
(e) The number of probesets peaking at different times of day. (f) A heat map showing the normalized levels of 1667 circadian transcripts in the SCN over two circadian cycles. (g) Normalized levels of probesets that belong to the gene ontology terms shown indicate that these gene families oscillate in coherence suggesting that these cellular functions are required at different times of day.
Figure 2.2. Light modulated transcription in the mouse SCN. 
(a) Tissue sampling times for the impact of light on the SCN transcriptome. A light pulse was delivered at CT30, CT40 or CT46 and SCN samples were collected 1, 2 and 4h after each of these three time points. (b) A heat map showing the transcript levels for 506 probesets that are affected by light. (c) qPCR validation of candidate genes. (d) A radar plot of the phase of probesets that are both rhythmic and light affected. Red indicates light suppressed and blue indicates light induced. (f) The overlap between light affected and rhythmic genes in the SCN.
Figure 2.2 (continued). Light modulated transcription in the mouse SCN. (e) The effect of light on transcript levels was examined for candidate genes in wildtype, melanopsin cell ablated retinas, mice that lack rods, cones and melanopsin RGCs, melanopsin knockout and rod-cone deficient mice.
Figure 2.3. SCN specific loss of Lhx1 disrupts circadian rhythm of locomotor activity.
(a) The flowchart of the comparative procedure used to identify SCN-specific genes. (b) A heat map depicting the expression patterns across 82 mouse tissues and the SCN for 230 probesets whose expression was found to be high in and specifically restricted to the SCN. (c) Actograms showing the locomotor activity patterns of indicated genotypes in 12h light: 12h dark (LD) or constant darkness.
Figure 2.3 (continued). SCN specific loss of Lhx1 disrupts circadian rhythms of locomotor activity.
(d) mRNA levels in the SCN of Lhx1 transcript in the SCN-specific Lhx1 deleted mouse. (g) Temporal patterns of mRNA expression for two circadian genes Per1 and Dbp in wildtype and Lhx1 SCN KO mice. (f) Transcript levels of genes involved either in the core clock machinery
Acknowledgements

Chapter 2 is being submitted for publication under the following title: Megumi Hatori, Shubhroz Gill and Satchidananda Panda, “Transcriptional landscape of the SCN and the role of Lhx1 as a determinant of SCN function”. The dissertation author was the co-first researcher and author of this paper, along with Dr. Megumi Hatori.
Chapter 3. An Infrastructural Setup for Studying Feeding Time in People

In the previous chapter, I described how light can affect gene expression in the mouse SCN. A key timing cue for the periphery is food. Food intake is itself governed by circadian timekeeping, and intake of food at specific times of the day then feeds back into the circadian clocks in the periphery thus providing a constant synchrony between metabolism and food availability in the environment. In this chapter, I will describe a methodology that I have developed to examine behaviors related to food and beverage consumption in people.

In peripheral organs such as the liver, delivering feeding-time cues even in the absence of a circadian clock can effect gene expression changes (Vollmers et al., 2009). This argues that there are circadian clock-independent connections between metabolic pathways and feeding time. As mentioned earlier, in mice under normal diets, the time of feeding itself is under circadian control, and then feeds back into the clock. It is possible that different diets that the organism has not evolved to consume disrupt the normal feeding/fasting cycle and thereby may impact the timing cues provided by feeding time to circadian metabolic pathways. Overall, since many metabolic pathways are localized to defined temporal niches, this can result in aberrant and interfering biochemical reactions occurring at the same time when they should not, or serial chemical reactions becoming delayed because an intermediate was not produced at the correct time.
Evolutionarily, organisms have developed sensory mechanisms to detect energy rich foods to psychophysically enhance their palatability and provide favorable stimulus through the taste sensing mechanisms. In the wild, this was an adaptation that was aimed at utilization of dense energy stores as and when they became available, and then potentially using them when food is rare. In the modern world, consumption of diets that contain extraordinarily high concentrations of energy has become the norm, at least in the developed world. There is usually no dearth of food at any point of time, and this increased consumption of calorie dense diets, such as high fat food, has been considered as the root cause of the obesity epidemic in the recent times. In experiments performed in our laboratory by my colleagues (Hatori et al., 2012), it was discovered that laboratory mice maintained on a high fat diet abolish the normal daily cycle of caloric intake, that is they eat at times when they would normally be fasting. To test the role of this aberrant feeding fasting cycle in causing obesity and its related symptoms, two sets of mice were maintained on a high fat diet. One set was provided access to the food whenever it wanted (ad libitum feeding) while the other group only had access to the food for 8h a day. Surprisingly, it was found that even though both the groups of mice consume the same number of calories from the same food source every day, the mice on ad libitum high fat diet became obese, while those on time-restricted feeding did not. Many pathological symptoms that accompany obesity such as inflammation, liver disease (hepatosteatosis) and diabetes are also absent in time-restricted feeding. This suggests that the disruption of the normal eating/fasting routine drives the
obesity outcome in mice maintained on a high fat diet. It is known in mice that the consumption of fatty food disrupts the feeding/fasting cycle, i.e., mice begin to eat at times they otherwise would not be eating. Feeding time and light are two of the important cues to organismal timekeeping mechanisms. This brings up the possibility that the adverse effects that are attributed exclusively to tipping the caloric equilibrium may also partly arise from a disruption of the organism’s timekeeping mechanisms.

The studies described above were performed in mice. We wondered what, if at all, significance this has for humans. It is well accepted that a person’s diet is a key determinant of his/her health. What has been lacking is an ability to make precise connections between aspects of the diet and their specific effects on heath. With the increasing adoption of fatty and sugary foods, individuals in developed countries are departing from the food they have evolved to consume. We are exiting not only our temporal ecological niche by remaining awake for long hours at night, but also our feeding time niche, by eating at night. The feeding patterns of modern day individuals are not only governed by hunger and satiety, but also social interactions, accessibility to food (such as vending machines) and how people divide their time between work and home. It is not surprising that the extent of obesity is mirrored in the timeline of industrial progress of nations. A third of the US population over 20 is now considered obese (Flegal et al., 2012). There is an urgent need to quantify the food and patterns of eating in modern societies. However, the advancement of nutritional
sciences in humans has been impaired by its reliance on poor quality assessment methodologies.

There are three methods that have been used in the past to assess feeding behaviors are:

Questionnaires: In this method, the person arrives in the lab and is asked to recollect what he or she ate or drank in the last 24 hours. This method relies heavily on a person’s ability to recall and a vast body of literature exists that shows that the correlation between the ground truth and responses to questionnaires is very poor.

In-lab monitoring: In this methodology, the subject is asked to stay in a closed setting with continuous monitoring, such as a video camera in a laboratory. The subject is presented a finite choice of foods. Notably, the subject is now constrained to choose from the provided food items and the convenience and cost factors that usually contribute to food choices are completely eliminated. Furthermore, the video monitoring and segregation into a novel environment itself is likely to alter behavior.

Food diaries: In this method, the person carries a pen and a notebook with him/her at all times. When the person is about to eat or drink something, s/he takes the pen and paper and writes down a description of the food. Since it
takes a significant amount of effort to write down, this is a highly inconvenient method for the subject. Additionally, at the researchers’ end, the scientists interested in studying the food choices would now receive a description of the food and thus there is a possibility that the description may be incomplete or inaccurate. Food diaries have been known to alter behavior because a person has the option to go through what he or she wrote. This feedback can alter behavior e.g. if a person sees that he/she has eaten two hamburgers already, he/she might refrain from eating another one after gaining this knowledge.

We thus sought a method that would allow us to examine these behaviors in people in the course of their normal daily lives without significantly altering their natural patterns. The method that we sought must be cheap, scalable to a large number of people and evidence-driven. We leverage the burgeoning penetration and dropping ownership cost of smartphones and present a workflow for studying natural behavior in the context of people’s usual daily life. We present a scalable methodology for collecting and analyzing feeding data in real-time from human populations, and apply it to study the extent of variation in feeding-time. Unlike previous approaches, our methodology is evidence-based and inherently time-stamped. In addition, by using geographical location to determine local sunrise and sunset time, we can appropriately connect feeding patterns to light availability.
In order to use a smartphone approach, I wrote a software application ("app") for Apple’s iOS platform in Objective C language. In its most essential incarnation, the app allows a person to take a picture and transmit it wirelessly to a remote server. The application is written in Objective-C and was initially distributed offline but is now available in Apple’s App Store.

The app has 4 tabs:

Pictures: In this tab the person can take a picture of the food or drink he or she is consuming. If the person is interested in describing the food/drink or the amount of the same, this can be conveniently done using another button. Once the person presses the send button, these two pieces of information (picture and its description) as well as the date and time and the location at which the picture was taken are sent to the server wirelessly over either the cellular internet or wifi. Study subjects are expected to use this tab for the most part. The picture is taken just before a person starts eating or drinking and is sent immediately.

Presets: If a person is at a social gathering, forgot to take a picture or any other circumstances prevent the collection of picture data, the person can choose from a list of common foods what he or she ate or drank and how long ago. These data are then sent along with location information to the server.
Manual Entry: This tab is very similar to the second tab, except that in this case the person can type out the name of the food or drink and an arbitrary time for how long it was consumed. Together, tabs 2 and 3 account for a person’s inability to remember to take pictures.

Info: In this tab, we provide some basic information about the study to the participant, as well as links to share information about the study with friends. The tab also contains a link by means of which a participant can email the researchers. We also remind participants that we collect location information every time they send any data to us.

To limit the usage of the app and transmission of data to our servers to participants in the study, we use a registration mechanism. When a person first installs the app, a unique user ID is created for him or her. Usage of the app is blocked until the person provides a unique access code. This access code is only provided to registered participants. Once a person enters a correct access code, the study identifier is received in a database on the server.

The server software that handles data uploads is coded in PHP. Databases are MySQL. The app names files in the following format: <userID>_<date>_<time> and this allows us to automate our analyses of the time of eating. Custom scripts written in PHP perform various analyses of the data upon demand.
For this study, we were interested in identifying the behavioral patterns as well as nutritional choices of healthy adult individuals. We defined the following recruitment criteria for our study:

1. The person should be an adult between the ages of 20 and 50 years.

2. The body mass index (BMI), defined as the body mass divided by the square of the person’s height in SI units, should be between 16 and 27. In the clinic, “normal” BMI is defined as 18-25. Thus we are allowing individuals in our study who are possibly slightly overweight or slightly underweight. We also recruited some obese individuals for the purposes of comparison.

3. The person should not be a smoker.

4. The person should have no prior diagnosis of metabolic or cardiovascular disease.

5. The individual should not have any eating disorder, bile syndrome or major immune disease.

6. No incident of sickness in the prior two weeks.

7. Female subjects should not be pregnant.

8. The person should not be enrolled in a weight management program or be taking any drug or medication that affects weight.

9. The subject cannot be an employee of the Salk Institute or an acquaintance of the researchers.

The workflow for recruitment of subjects for the study and data collection is as follows. Advertisements for the study are posted both online and offline. In
response to the advertisement, a person goes to our website and fills out information about himself or herself. These information are name, email address, age, height, weight, ethnicity, location, whether the person is a smoker, has any major disease(s) and if female, whether she is pregnant. We also ask the participants some simple questions about iPhone usage to ensure he/she is familiar with most of the common functions of the same.

Once the person provides this information, we screen out individuals who do not meet the inclusion criteria. After this screening the person is asked to come to the laboratory for the first visit. During this first visit, we measure his/her height and mass to verify the self reported values. We then administer a standard questionnaire about the person’s eating habits and chronobiological patterns. The person then signs the consent document. The first height and weight measurement is performed within one week of the start of the data collection. The data collection begins at midnight on a Wednesday and goes for three weeks. After the completion of the three weeks of the data collection, the subject returns to the laboratory for a second visit, which is scheduled within one week of finishing. In this visit, we measure height and body mass again, and compensate him or her for participating.

The participants are told that the goal of the study is to estimate the diversity of the kinds of food and drink that people consume, and that our long term goal is to understand the effect of specific diets and dietary components to
specific aspects of health. It is not disclosed that we are interested in the time of eating or drinking.

They are instructed to always carry their phone, to take a picture of everything they eat or drink regardless of how big or small it is and to send the picture immediately. Next, they are informed about the compensation and duration of the study, and the voluntary nature of the study i.e. they can leave the study whenever they want.

One of the features of the smartphone application is that the participant can annotate information about the food or the quantity in case he or she thinks we will be unlikely to identify what it is or how much. Examples of some use cases for this are: the kind of milk in coffee can be described, if a person is eating a sugarless cookie, she can describe that or if there are multiple persons with whom the item is being shared, that can be described.

To protect participant privacy, identifiable information is not published or presented. Geographical location data is scrambled into a referential form rather than absolute latitude and longitude coordinates. Finally the subjects are informed that any personally identifiable information in the picture itself, such as a person’s face, an ID card are all removed by us.
If, for example, the person is in a basement and there is no internet, the app displays a warning and the person must resend the picture from another location. If the subject visits a location where there is no internet such as during a hike, he/she must continue taking pictures with the camera and email the pictures and times to us later on.

Water, coffee, medicines, alcohol and chewing gum are all regarded as food for the purpose of this study, and the participant must take a picture of it. In case a person is eating or drinking slowly, and the time difference between two consecutive sips or bites exceeds 15 minutes, they are instructed to regard it as a separate eating event and take another picture, as well as describe it as being the same item. Participants are reminded to take pictures of anything that they eat or drink late at night or early in the morning, and that for this reason, they must keep their phone close to their bed at night. If a person sends a picture of a food item but does not consume all of it, he or she is asked to send another picture with the remainder and use the annotation button to inform us that this is the same item as earlier.

A subset of the participants is requested to wear a MotionWatch device. This lightweight waterproof device has a triaxial accelerometer that detects any motion along any of the three (x, y or z) axes. The MotionWatch is worn on the wrist like a wristwatch. The MotionWatch also detects ambient light levels. Those participants who are asked to wear a MotionWatch are requested to wear it for all
21 days of the study. The collected data are transferred at the end of the study to a computer and analyzed either by the specialized software provided with the actigraph itself or can be exported as raw data for analysis in other programs. Actigraphy is used to determine various parameters concerning a person’s sleep/wake cycle and activity levels. Examples of parameters that can be extracted out of long-term actigraphy include sleep duration, sleep latency, mean activity, fragmentation index and number of sleep bouts.

To assess the reliability of our method, we developed another independent data collection pipeline. In this method, about twice a day, at a random time of the day, we trigger a push notification to the subject’s iPhone. This results in a message being displayed on the subject’s smartphone immediately. If the subject sees the question at 11:34am, for example, the question is “Did you eat or drink between 10am and 11am today?” That is, the question always pertains to the last complete hour. This allows us to minimize reliance on memory recall since in the worst-case scenario (if the person sees the message at 59 minutes past the hour), the subject has to recall information about the last 2 hours. The person responds with a yes or no. The subject’s response, together with the presented question, and the time at which was presented is recorded in our database.

The content of the question depends on when the person saw the notification and not when it was triggered from the server. This is to allow for the
possibility that the subject may not notice the notification for a while after it has been sent, and yet the question does not exceed the intended time duration into the past. If the person sees the question at 10:43am e.g., the presented question is “Did you eat or drink between 9am and 10am today?” Thus, in the best case, the person is required to recollect for the last 1h and worst case, for the last 2 hours. By avoiding memory recall any farther into the past than two hours, we are hoping to avoid the pitfalls of 24h questionnaires.

We then compare the answer to the picture data for the same time interval. There are four combinatorial possibilities (response/picture data):

(a) yes/yes,
(b) yes/no,
(c) no/yes and
(d) no/no.

Combinations (b) and (c) define the type I and II errors respectively, i.e. false positive and false negative.

Given the nature of the data being collected, we make a reasonable a priori assumption that the distribution of eating or drinking events in time is sparse, that is a person eats for a small fraction of the time while most of the time is spent not eating. This is a very valid assumption for nighttime but even in daytime this is in fact reasonable since even if a person eats 20 times a day and
spends 15 minutes (very lax assumptions) on each of those, the total time spent on eating would still be 5h, or less than a third of a day. In my experience, the total time spent eating on average is actually closer to 2h. Thus, for a push notification sent at a random time of the day, the response is more likely to be no than yes.

I also designed an automated system to keep track of data as it arrives. Every half an hour, the system analyses the data and identifies individuals who have not contributed any picture data in the last 8h and sends this information to the researchers. Using this system, we can keep track and identify any individuals who are not thorough with data collection. In addition, a daily, automated email summarizes all the data gathered within the last day.
Figure 3.1. Interface design for the user-facing smartphone application.
This figure depicts the user-facing screens of the smartphone application used for the study. The app consists of four tabs. (a) The first tab is used to take a picture of a food or drink that the person consumes and to send it wirelessly to our server, using either cellular internet or wifi. (b) If the user deems necessary, an annotation describing the food itself or its quantity can be provided. (c) In case a user forgets or cannot take a picture of a food item, he/she can choose from a list of common foods or (d) type out what he or she ate. Finally, the information tab shown in (e) provides some basic information about the study and a means to email the researchers if the participant has questions.
Figure 3.2. Nature of the data received for each food or drink instance.
Every time the user takes a picture of what he or she ate or drank (or chooses from presets or types out), we receive three separate files. The first is either a JPEG image that the user captured, or a text file with the extension .TXT in case the person used a text option for describing the food. The second is a text file with the extension .DSC that contains the optional annotation for the image provided by the user. Finally, a third text file with the extension .ASC contains the latitude, longitude and the GPS resolution for the place from where the picture was sent. The naming convention for each of these files is <user ID>_<date>_<time>.
Figure 3.3. Study workflow for subjects and researchers.
Figure 3.4. Push notification system.

An orthogonal approach was implemented to assess the reliability of the user-initiated data gathering. This system makes use of Apple’s push notification system that can be used to deliver small textual messages to a user’s smartphone. The flowchart above describes the sequence of events. Approximately twice a day, at a random time of the day, the server triggers the push notification. Instantaneously an alert appears on the user’s smartphone prompting him/her to answer a question regarding the study. The times T0, T1 and T2 at which these events occur are almost the same. At a time T3, which may be long after T2, when the person notices the question and intends to answer it, a question pertaining to the last complete hour from T3 (and not T2) is presented. The user responds to this with a yes or no. Instantaneously, the user’s response as well as the question presented and the time at which it was shown is recorded into a central database on the Salk Institute server.
Figure 3.5. Push notification system: user-facing screens.
The above screenshots show (a) the notification of a new question and, (b) actual question presented to a user. Note that since the time at which the person saw the question is 7:42 pm, the content of the question pertains to 6 pm to 7 pm.
Acknowledgements

Chapter 3 will be submitted for publication and the dissertation author was the primary researcher and author of this paper.
Chapter 4. Insights from an Evidence Driven Approach to Studying Human Behavior

I will first describe the characteristics of the demographic studied. In most human behavior studies conducted in academia, it is customary to use university students, especially undergraduates, as subjects. (Henrich et al., 2010) This has led to the idea that most of the data concerning behavior gathered in humans is in fact data that pertains to college students. Although this was the first human study conducted by our research group, or myself we attempted to bypass this drawback of human behavioral research. We advertised to the general public at large to attract them to participate in our study. Furthermore, by the end of the study, we intend to recruit individuals in the age range 20-50 years. Also, we attempted to enroll equal number of males and females, and have no preference for or against any ethnicity. In this way, we are hoping to have a much more realistic and general picture of what ordinary people eat and drink, and the patterns surrounding the same. We did not select for or against any chronotypes either and thus we expected to find some “larks” (early risers) and some “owls” (late risers).

This dissertation describes results from 64 study participants. The age range is between 20 and 44 years. The study population is relatively diverse in terms of age and body types as can be seen from figure 1. We have also been
able to sample a sizeable fraction of the ethnic diversity present in our study location (San Diego) as shown in figure 4.2.

For every behavioral assessment methodology, it is vital that the act of observation should not alter the subject’s behavior. In our case, we do not wish to alter the participants’ eating behavior. To assess the impact of participating in our study on the physiology, we measured the height and body mass of the subjects before and after the three weeks of the study. To keep the measurements accurate, both sets of data were gathered within one week of starting or finishing. Figure 4.3 shows the change in BMI for all 64 participants. As can be seen from the figure, it did not change. The median and ranges of the variables are not different.

In figure 4.4, we assessed whether the minor changes observed in BMI were in any way related to the person’s initial BMI. That is, did (for example) heavier people lose more weight? This figure shows that the BMI change is not correlated with the initial BMI.

After obtaining the data, we examined the temporal ingestion patterns. These are plotted in figure 4.6 where we see that most individuals do not have a defined pattern of food intake. There is also a surprisingly large incidence of late night eating (i.e. after 12 midnight). We then determined the high confidence interval for these data points which is shown in figure 4.7.
Because of the emergence of sedentary lifestyles, people often resort to keeping bowls or packets of food or drink close to their person, and eat or drink continuously at short intervals. To identify the nature of such continuous eating, we analyzed the distribution of time difference between consecutive eating or drinking events, as shown in figure 4.8. It can be seen that most people eat at relatively short intervals of 2-4h and this counters the “three meals a day” myth. Examination of the median of all intermeal times for different individuals one by one shows that these are centered around 1-2h, as shown in figure 4.9.

We developed a method to present the data in a manner that is easy to visualize and yet conveys several layers of information coherently at the same time. We call this a feedogram, as shown in figure 4.10. In this plot, all data received are binned into 15-minute intervals. Each row is one day. The data from left to right are plotted from midnight to midnight. Time between local sunrise and local sunset time is regarded as daytime and shaded yellow, while the remainder of the 24 hours is shaded gray and regarded as night time. This allows for a convenient visualization of evening and early night eating on the right hand side gray area. Late night eating can be seen in glance in the left hand side gray zone. By clicking on any of the vertical line marks, the web interface takes the user to the actual picture collected during that time interval.

The feedogram also makes it very easy to identify any instances of multiple eating events spaced closely apart, which represent consumption binges.
The onset of feeding as well as the time at which it ends can be easily seen too. Finally, any additionally layers of information such as ambulatory activity patterns can be easily added either as a row below the food data row or as background color.

While picture data is information and provides the best form of evidence for the eating behavior, it provides a computational hurdle. Automated analysis of the contents of a picture is not a viable computing strategy. Thus, to provide semantic structure to the pictorial data, we manually annotated the pictures and the results of the same are presented in figure 4.11. In the process of annotating the pictures, we also estimated the caloric content and the volume in ml of the items consumed. This allows us to address when do people consume their largest meals, both in terms of volume and calories. These are shown in figures 4.12, 4.13 and 4.14.

As described previously, our experimental strategy incorporates push notification based messaging to gather data on feeding times independently of the user-contributed pictures. By comparing the results from these two methods, we can assess the reliability of our experimental approach (figure 4.15). If we compare the push-messaging answers with the same time interval the true negative rate is high while the true positives are low. This is because people do not remember the exact time at which they ate, even in the short term span of
about 2h. The true positive rate improves if we expand the interval to 1h on both sides.

The feeding time onset and offset indicated that the duration of feeding might be closely tied to the persons’ waking period. Thus we measured ambulatory activity patterns by means of wrist actigraphy. The results (figure 4.16) indicate that there is a huge variation in the total activity levels between individuals. Our results in figure 4.17 also show that most individuals stay up until late night every day. Such insights were not possible in the past because of the absence of actigraphy systems that could be sustained for weeks. The MotionWatch 8 system that we have employed allows us to track behavior for upto 3 weeks, and thus our work reveals previously unknown and underappreciated aspects of diurnal behavior of modern day people.
Figure 4.1. Age and BMI distribution of the study population.

BMI (body mass index) is defined as the ratio of the body mass to the square of the height of a person in SI units. Normal BMI is defined as 18-25. For this study, aimed to recruit individuals in the BMI range 16-27, but as shown above, some of our participants were overweight (BMI>25) or obese (BMI>30). More study participants belong to the 20-30 age group, than 30-40 or 40-50.
Figure 4.2. Self-described ethnicity of the study population.
Participants were asked about their ethnicity when they first filled out an online form expressing interest in participating. Majority of our participants are Caucasian, with Asians (including those from the Indian subcontinent) making up the second largest group. About one-eighth of the participants were of mixed descent. About 1 in 10 were of Hispanic origin. The smallest category was African-Americans who comprised about 1.5% of the study sample.
Figure 4.3. Initial and final BMI of the study population.
We measured the height and weight within 7 days of starting and used that to evaluate the initial BMI. These were measured again at within one week of finishing the study for final BMI. The values are plotted above.
Figure 4.4. BMI change versus initial BMI.
In this graph, the change in a given subject’s BMI is plotted as a function of the initial BMI. Most points are confined to the ±0.5 along the y-axis, and all points are confined within ±1. This indicates that our subjects BMI did not change substantially during the study. Furthermore, the $R^2$ value is approximately equal to zero, indicating that the change in BMI did not correlate with the initial BMI.
Figure 4.5. Example of pictures received from one person in one day. These eight pictures were received from one single person in the course of one day.
Figure 4.6. Distribution of the time of eating or drinking for different individuals.
Every dot in this graph depicts the time at which a person sent a picture, and thus, consumed a food or a drink item. To maintain continuity in the graph, any points between 12 midnight and 4am are shown as 24h+the time at which they were sent. As a result, the y-axis is from 4h to 28h.
Figure 4.7. 95% interval of the time of eating or drinking for different individuals.

(a) The bar graphs represent the 95% interval for the points in figure 4.6. The sequence from left to right is the same as in figure 4.6. (b) In this case, the bar graphs of figure 4.7(a) are arranged according to the start time.
Figure 4.8. Inter-meal time for different individuals.
This graph shows the distribution of the inter-meal time, defined as the time difference between two consecutive pictures received from the same person. The data is truncated at 24h, i.e. any inter-meal times greater than 24h are not plotted in this graph.
Figure 4.9. Median intermeal times for individuals.
The median intermeal time (MIT) for an individual provides an estimate of how frequently a given person eats. This graph shows a histogram for the MITs for different individuals studied.
Figure 4.10. Feedogram.
We have devised a novel visualization for presenting the times at which the person consumed an item. This is called a feedogram and is shown in the figure above for one individual. All received pictures are binned into 15 minutes intervals. A dot represents the absence of any eating or drinking event in a given 15 minute interval, which a vertical line represents the presence of the same. This visualization is generated automatically in a web-based interface by a script. Clicking a vertical bar shows the data collected in that 15 minute interval. The background color is gray or yellow depending on whether it is daytime (sunrise to sunset) or nighttime (sunset to sunrise). Late night eating (after midnight) can be conveniently seen on the left hand side gray area.
Figure 4.11. Annotations reveal patterns of food and beverage consumption. Total events captured were annotated for different parameters. (a) Physical state of the item consumed. (b) Temperature of the item consumed. (c) Taste modality activated by the item consumed. (d) Does the item consumed contain alcohol? (e) Does the item consumed contain preservatives? (f) Is the item ready to eat?
Figure 4.12. Estimating the calories and volume of each item consumed.
For each item captured in a picture, we estimated the calories and volume for it. Those are plotted in this graph. Most items are less than 1 liter volume and 1500 calories.
Figure 4.13. Temporal spread of caloric intake.
Each dot represents one picture. Along the x-axis is the local time of day at which the picture was taken. It can be seen that most of the calories are consumed in lunch and dinner at around noon and 10 pm respectively.
Figure 4.14. Temporal spread of volumetric intake.
Each dot represents one picture. Along the x-axis is the local time of day at which the picture was taken. Ingestion of food or drink affects the stretch reflex in the stomach which signals satiety. Thus, it is important to not only consider calories, but also the volume of the item as shown here. Again, it can be seen that the major meals in terms of volume are lunch around noon and dinner around 10pm.
Figure 4.15. Estimation of specificity and sensitivity of the methodology. Answers provided in response to push notifications sent out at random time of the day (during the average person’s waking hours) were compared to whether or not one or more pictures were received in the same time interval for the top grid. For the bottom grid, the time interval for the pictures was extended by 1h on each side.
Figure 4.16. Relative activity levels for 20 individuals.
We assessed the activity levels in study subjects using a MotionWatch. The counts for the MotionWatch actigraph reveal the number of times the arm on which it is worn moves in any of the three axes. For this graph, the total counts in the 21 days were normalized to the maximum value and plotted.
Figure 4.17. Temporal spread of average activity levels.
MotionWatch counts for a given person in each hour of the day were averaged for the 21 days, and plotted to examine the times at which a given person is, for example, most and least active. Along the y-axis, the hours of day are plotted from 1 (12 midnight to 1am) to 24 (11pm to 12 midnight). The color intensity is proportional to the average counts.
Chapter 5. Future Directions

Our survey of nutritional choices and behaviors in individuals who have a relatively normal body weight provides a baseline for further studies. For overweight or obese persons, our mouse study hints that time-restricted feeding can be used to prevent metabolic disease and obesity if these individuals eat high fat diets for extended hours. Thus, in a subsequent study is the future direction for this project, we are recruiting obese individuals (BMI of at least 26) who go through a round of assessment using the iPhone app for three weeks. They are told that their patterns and nutritional choices will be analyzed to determine if they will be eligible for the intervention stage.

Participants from our previous studies who have informally expressed interest in future studies are also a potential pool for the intervention, and they will not undergo the preliminary data collection. In these three weeks, they are asked to take pictures of everything they eat or drink, as well as requested to wear a MotionWatch. From the pool of overweight applicants, we identify individuals who consistently eat for more than 16h a day. These individuals will be asked to reduce the number of hours that they eat to 12h or less per day. There will be absolutely no restriction on what they eat in those 12h. In the remaining 12h, they are only permitted to drink water. This regimen is carried out at their home or workplace in the course of their normal daily life and not in a laboratory setting.
The 3 weeks of data collected from such individuals is regarded as the preliminary data, and is the basis for advancing a given individual into the next stage.

Identification of individuals whose eating patterns might benefit from the time-restricted feeding regimen:

From the 3-week long preliminary data collection, we will shortlist individuals who meet the following criteria as prospective subjects for intervention:

1. BMI of at least 26.
2. Starts eating between 7am and 10:30am
3. Eats for at least 15h every day, consistently.
4. Responds to all push notifications in a timely manner.
5. Does not remove the MotionWatch activity monitor for the 3 week duration of preliminary data collection, if she/he had agreed and was requested to wear one.
6. No travel outside of PST time zone for the study duration.
7. Indicates that he/she does not weigh himself/herself regularly and a verbal commitment to not weigh himself/herself for the study duration.
8. Has not participated in a clinical study in the last one year. This is to avoid confounding effects from previous therapies/placebo the person may have undergone.
9. No surgical intervention for weight management.
Individuals who are not identified as suitable for intervention are asked to come to the laboratory again for a second height and weight measurement, and to collect their compensation. After a person has been shortlisted as being suitable for intervention by the above criteria, we will email them with an offer to participate in the intervention study. The email will explain the scientific rationale for the intervention, the potential benefits to the participant, compensation details and a description of the intervention itself. If the subject is interested, he/she is asked to visit the lab again where we will show them a presentation about the details of the intervention and have them sign the new consent form for the intervention if they are interested in participating.

Participants are informed that their eligibility for the intervention is subject to he/she not being hypoglycemic. During the visits prior to and after the intervention, the shortlisted subject is asked to assess her wellness using a questionnaire. The intervention subject’s height is measured at the intervention onset. Weight and body composition are measured using a Tanita BC-1000 scale. These are measured again at the end of the 16-week intervention. There is one long-term measurement of all these parameters once 1 year after the intervention start. Blood samples will be collected by a clinical facility after overnight fasting. For every shortlisted participant, there will be an initial blood sample collection, a portion of which will be used to perform the Comprehensive-7 blood panel. If the person is hypoglycemic as determined by the blood panel report, he/she is told that he/she is ineligible to participate. For subjects who undergo the intervention,
there is one blood draw after 16 weeks (at the end of the intervention), and another long-term blood draw 1 year after the intervention start.

From the preliminary data and in consultation with the individual subject (to take into account his/her work and personal life schedules), we will determine a suitable daily eating start time (DEST). The subject should not eat before DEST. Daily eating end time (DEET) is 10-12h after DEST. No conditions are imposed on what the person eat or drinks during the permitted hours. The participant is advised to stop eating at DEET. After a given day's DEET and before next day's DEST, the participant is allowed to drink water.

We have developed a novel visualization of a person’s daily eating pattern that summarizes multiple days or weeks of eating behavior in a nutshell. We term this diagram a “feedogram”, and is our first piece of feedback. In one glance, it shows the following:

1. The temporal spread of eating or drinking events.
2. The phase lag between eating start time on weekends and weekdays.

The second visual feedback is the distribution of the time interval between two consecutive meals. We term this the IMT (inter-meal time) diagram.

The third piece of feedback is a collage/collection of all the pictures obtained from a particular person for the last day and for the last week. These
are presented together or one by one, depending on whether the pictures are high resolution or not. The iPhone model determines this.

The fourth feedback item is a tag-cloud for all items in the prior 7 days.

These will be sent to the subject via email in the following manner. The subject is sent a hyperlink to a customized page on metabolicstudy.salk.edu, where the feedback items are presented one after the other on different pages. One feedback page will lead to the next one. By directing the person to a website, we can ensure that all pieces of feedback are viewed one by one. At the end of the feedback, there will be a short question about the contents of the feedback items. The question is formulated so that it is trivial for a person who has paid attention to every feedback item, but cannot be answered by a person who has not.

One day prior to the start of the intervention, and one day after the end of the intervention, up to 6 ml of blood will be collected. The blood collected prior to the study will also be used to perform comprehensive panel and an on-the-spot fasting glucose test to ascertain if the person is diabetic.

One portion of the blood will be sent for metabolite analysis. The remainder will be retained in case the first sample fails metabolomics quality control (QC) checks. No DNA genotyping is performed on the blood. We wish to
shortlist metabolites that have previously been annotated as detrimental to health and the expectation is that these would change in the direction and to levels closer to those found in healthy people.

This intervention will last 16 weeks. We will compare the body weight and other physiological characteristics in these individuals before and after the intervention to address whether temporally restricted feeding improves any health parameters. In particular, we will examine the body mass, body composition and blood metabolites. Blood will be collected from these individuals before and after and sent for quantitative metabolomic analysis using GC/MS. The goal of the blood metabolomics is to identify any metabolite(s) that are elevated in obese individuals relative to healthy people and is/are improved in a majority of the obese individuals. Improvement here would be defined as a change in the direction that is found in normal individuals. Our laboratory has previously performed GC/MS metabolomics using mouse tissues and in those experiments, more than 300 metabolites were successfully quantified. Thus I am optimistic that the metabolomics will provide insights into the health effects of time-restricted eating in people.

The second future goal of the study is to expand the data collection to a much larger population. This can be easily done by removing the blockade at the first launch of the app. Any one would be able to contribute to the data collection anonymously. The goal of this study would be to collect data from thousands of
individuals all over the United States and possibly, other parts of the world. By collecting this data, it would be possible to compare the dietary patterns in different geographical regions and potentially correlate them with disease incidence to discern meaning patterns and generate hypotheses that can then be directly tested. Furthermore, the duration of data collection can be increased to several months or possibly even multiple years. As a result, it would then be possible to examine how a person’s eating patterns change with changing seasons. It can also be determined in a quantitative manner how patterns change during holidays such as Christmas or Thanksgiving when it is anecdotally known that the consumption of sugary foods spikes. A comparison of the same person’s data before, during and after such salient days would describe the impact of people's social engagements on their health. The main disadvantage would be that the individual’s height and body mass cannot be verified.
References


