“Ome” sweet “ome”: From the Genome to the Conductome

Christopher R. Stephens

Abstract The last few decades have seen science both changed and confronted by the appearance of enormous quantities of data, that have arisen from the development of multiple new technologies. The impact of this “data revolution” has been particularly acute in the biological sciences, where bioinformatics has made great strides in integrating such data into new theoretical frameworks and adopting new computational tools. One framework that has prospered is that of the “ome”, which adopts a more holistic view of the physical structures that make up a cell, tissue or organism and their mutual interactions. The structures associated with the principal “omes” - genome, proteome, transcriptome and metabolome - are all microscopic, being associated with different biological molecules. Recently, however, the omic approach has been applied to more “mesoscopic” structures, such as organs and tissues, with the resulting totality of structures conforming the physiolome. However, all these omes are associated with particular spatial and temporal scales, and are therefore inadequate for addressing the real complexity of living systems, which are both multi-scale and highly multi-factorial with respect to those scales. We argue that a “disease-ome”, for example, as the totality of factors associated with a given disease, requires the integration all the current omes, and more. Thus, a holistic description of an important disease, such as obesity, requires all micro, meso and macro factors, as well as an understanding of both their upstream and downstream causal relations. This is particularly challenging when the relevant factors are distant in scale. Thus, the causality between overeating and obesity at the individual level is clear. However, the link between a certain genotype and obesity or the link between food production and obesity is much less clear. In spite of this, all of these factors can, in principle, be collected and included in a prediction model,

Christopher R. Stephens
C3 - Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Mexico, CDMX 04510
Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Mexico, CDMX 04510
e-mail: stephens@nucleares.unam.mx

using present technology and computational tools. We argue that the fundamental concept that most naturally links the micro, meso and macro is that of behaviour, as it is influenced by both micro (nature) and macro (nurture) factors and, in turn, influences them. We discuss the concept of the Conductome - the totality of factors that influence behaviour, using as an example food consumption and obesity, and emphasise its utility as an unifying concept that allows for a truly systemic view of a living organism.

Key words: Conductome, behaviour, genome, physiolome, complexity, obesity

1 The Micro-omes

In the last few decades there has been a trend towards a more global, systemic approach to the study of biological systems. A manifestation of this more holistic approach is the proliferation of different "-omes", and their corresponding fields - "-omics" [1], where the emphasis is on identifying as complete a set of relevant "structures" as possible that belong to the corresponding "ome". The original "omes", such as the genome, originated in molecular biology, where the relevant structures, such as genes, proteins, RNA transcripts and metabolites, are all molecules, and as such we can think of the corresponding omes as all being micro-omes. In the context of these micro-omes, the natural mathematical framework for understanding structure and interactions has been that of a network, where the nodes are structures and the links represent interactions. The information needed to construct such an "ome", as representing a totality of structures, is enormous, and has required the development of advanced technologies, such as high-throughput sequencing and capillary electrophoresis mass spectrometry in the case of bioinformatics [2].

Through a physicist's eyes, this "omic" approach seems no different in spirit to the traditional approaches to be found in physics and chemistry. The Periodic Table is, in that sense, the "atom-ome", the complete set of relevant structures at the atomic level. There is also an "elementary particle-ome" and a "moleculome". Indeed, the "omes" of molecular biology should be subsets of this moleculome. Although it is important to be able to identify the set of relevant structures, or the "building blocks", that form an "ome", this difficulty pales in comparison to identifying the complete set of interactions between these structures. The Periodic Table gives us the totality of atomic structures but certainly not the totality of atomic interactions. Thus, the set of possible molecules is much larger than the set of possible atoms. Additionally, interactions are manifestly contextual, meaning that the interaction between two structures is intimately dependent on the context of the environment they are in. For example, the interactions between two carbon atoms are quite different if they are in different DNA molecules versus the same one, or close versus distant in the DNA molecule, or within a chromatin complex or not.

Another way to put this is that: understanding interactions at one scale is no guarantee that we can understand interactions at another. Thus, just as an understanding

of atomic physics does not a priori allow for a quantitative, predictive understanding of molecular physics, as molecules are emergent structures relative to atoms, so an understanding of the genome "gene by gene" does not a priori allow for a quantitative, predictive understanding of what the genome does or, indeed, even what a small set of genes do, due to the presence of genetic interactions (epistasis) [4]. The difficulty of relating structures at one scale to emergent structures at another is one of the most difficult problems in science. Exceedingly difficult in physics, while in the biological sciences it is almost overwhelming. In physics there are several theoretical frameworks that encompass passing from the micro to the macro, the link between statistical mechanics and thermodynamics being the most developed. At a more general level, synergistics [3] has been applied to both physical and biological systems and attempts to delineate generic features and general principles, such as self-organisation and the existence of a relatively small number of order parameters, that lead to a description of the macro from the micro. The standard omic approach, in contrast, is more directly phenomenological in nature. Moreover, the current molecular omes are also associated with a particular type of structure: genes with the genome and proteins with the proteome. However, genes and proteins also interact with each other in a complex fashion, from the production of proteins by the transcription of DNA through proteins as transcription factors that control gene expression.

2 Disease-omes: relating the micro to the macro

Although the "omic" approach has its origins in the micro, recently the idea has been extended to more macro "omes", such as the physiolome [5, 6, 7, 8], where the relevant structures, such as organs or tissues, are much fewer in number. In this case, a network-based approach can naturally be applied and different interaction measures introduced, such as the degree of correlation in the time series of the different organs such as lung and heart. Like the genome however, the physiolome is associated with particular spatial and temporal scales that stem from the physical structures it considers. Thus, although a goal of the omic approach is to be less reductionist, the current omes are all very much linked to a certain scale. However, a true hallmark of complexity - of living systems - is its multi-scale nature, with relevant structures at many different scales. Thus, for instance, the heart, thought of as a physiological unit, has multiple associated spatial and temporal scales: the cellular scale, where pacemaker cells set the underlying heart rate, to circadian variations in the functioning of the heart at the cellular level [9], and on to the long term irreversible changes across a lifetime that are associated with heart disease. In this case, what we do at the macro scale, such as eating a lot of high-cholesterol foods, has an effect at the cellular level, leading, for example, to atherosclerosis which, in turn, has an effect at the macro level, where an artery becomes blocked, leading to a myocardial infarction. However, myocardial infarction itself has also been confirmed to have a genetic component [10], which then introduces a scale below the cellular

level, that of a single nucleotide. Unfortunately, as with many genetic studies, the causal chain that links the observed correlation between a micro property, such as a particular SIngle-nucleotide polymorphism (SNP), and a macro property, such as atherosclerosis and a subsequent myocardial infarction, is very poorly understood. Of course, the cellular level must enter as a relevant scale that links the two.

Thus, we are faced with one of the principal challenges of describing truly complex phenomena: the ability to incorporate structures, and their interactions, that exist at multiple scales. In other words, a micro-level molecular "ome" is in no way sufficient to encompass an important phenomenon such as atherosclerosis. However, neither is a "meso" ome, such as the physiolome, due to the need to incorporate micro factors such as SNPs. The challenge does not stop there, however. If we consider long term changes in heart structure and function, then two other principal categories of factor enter: aging and "lifestyle", which are the remit of macro-disciplines, such as epidemiology. Studying the disease can be done "bottom up" - trying to link macro effects, such as the clinical manifestations of disease, to the micro [11]. Indeed, much of the focus of the micro-omes has been to link macroscopic disease states to genomic, proteomic, transcriptomic and metabolomic data. In this case, the abstract mathematical framework is that of a conditional probability, $P(disease\ state|state\ of\ the\ genome,\ proteome,\ metabolome\ etc)$. The more conventional approach to disease prediction however has been "top down", linking the disease state to macro-variables such as age, sex, socio-economic status etc. A disease state however, is a complex multi-scale phenomenon, requiring a unification of the bottom-up and top down approaches, as in biological systems the micro and macro are linked and influence each other in a much deeper way than in physical systems. For instance, aging sounds simple enough to account for, but, as is known, chronological age and biological age are not the same, with the latter also having a genetic component. They are also linked by lifestyle, by which we mean the universe of external factors that affect the organism at multiple scales, from the genetic via, for example, environment-induced mutations, to the truly macro, such as the degree to which the environment itself favours the development of heart disease through, for example, diet. In contrast, in physics, with a nucleome, atomome, moleculome etc. we don't need the nucleome to understand the atomome. "Atomomics" can be developed in terms of structures - atoms - and their interactions without reference to the nucleome and its constituents - nucleons - and their interactions - the strong nuclear force.

We see then that if we wish to understand a phenomenon, such as heart disease, "omically", i.e., in the sense of a more holistic, non-reductionist perspective, it is necessary to go beyond an "ome" that is linked to a particular range of scales, as each only offers at most a partial view of the phenomenon. So, should we introduce the concept of a "myocardial infarction-ome"? where it comprehends all the factors that influence that outcome? This potentially involves the genome, proteome, metabolome, physiolome and several other omes that are still to be characterized, such as a "sociolome" or a "psycholome". However, at the same time the genome is linked to many more macro phenomenon than that of a myocardial infarction. Many, if not all diseases, have a genetic component. To try and be more precise:

imagine a set of diseases, $(C_1, C_2, \ldots, C_m)$, and a set of factors, $(X_1, X_2, \ldots, X_N)$, that are potentially related to those diseases. These factors include genetic factors, epigenetic factors, physiological factors, social factors etc. However, considering the effect of all possible causative factors on the set of all known diseases is not a recipe for success. We can group the totality of factors in a particular disease-ome in groups - genes/SNPs, proteins, metabolites, cell structures, tissue and organ changes, lifestyle factors etc. In this way however, we are faced with the perennial question of causation versus correlation. Is a SNP a "direct" cause of a disease or a correlative, indirect risk factor? Is socio-economic status a direct causal factor in the development of heart disease or a proxy for, potentially, many other more directly causal factors? What about diet? More directly causal? Then we have the fact that the impact of diet is influenced by the genome - nutrigenetics - while diet, in turn, affects gene expression - nutrigenomics. Life is complex. Literally. Nature affects nurture and nurture affects nature.

We argue then that the standard "omic" approach is still too reductionist to comprehend a complex phenomenon such as a disease, which is associated with structures and their interactions at multiple scales and where the interactions can be between structures that are naturally described at quite different scales. Although a network-based approach can be applied at the level of the disease-ome, by considering multiple diseases, for example, the natural framework for a given disease is, again, a conditional probability: $P(C = disease|\mathbf{X} = disease\ causes\ and\ risk\ factors)$ where, for example, C could be a disease state and $\mathbf{X}$ the set of factors that we wish to consider as conditioning factors on the probability to be in the disease state. Naturally, the data requirements to construct $P(C|\mathbf{X})$ are far greater than those of the micro-omes, where the latter are just one component of the disease-ome and, generically, not even the most predictive part. Furthermore, due to the multi-scale nature of the disease-ome, its construction through data has to transcend the disciplinarity that still exists as the principle foundation of scientific research.

Mathematically speaking, the disease-ome, $P(C|\mathbf{X})$, is a prediction model. Such a prediction model may be transverse or longitudinal, depending on whether or not C and $\mathbf{X}$ can be identified as states in time. For example, that C represents the development of a disease in a certain time interval and $\mathbf{X}$ represents the set of predictor variables identified in that time interval (transverse), or that they represent histories up to the beginning of that time interval (longitudinal). Of course, when $\mathbf{X}$ is high dimensional, a direct estimate of $P(C|\mathbf{X})$ is impossible, as $P(C|\mathbf{X}) = 0,\ 1$; i.e., every element is unique and either exists in a single element or doesn't. For example, no two genomes are completely identical, and the vast majority of potential genetic sequences of length ℓ - 4^ℓ - have never existed and probably never will. Thus, $P(C|\mathbf{X})$ must be estimated indirectly. There are, for instance, many machine learning based methodologies that can help in this regard. Seen abstractly, $P(C|\mathbf{X})$ represents a Bayesian belief network, where, in principle, if one could deduce its structure as a directed acylic graph would reveal the probabilistic relations between the different variables X_i, both among themselves and with the disease state itself. The goal would be to determine that graph that is most in accord with data. Unfortunately, computationally, this is an NP hard problem. Rather than search through a large

space of potential graphs, an alternative is to restrict the topology of the graphs. A particularly useful approximation in this regard is the Naive Bayes approximation, or generalisations thereof [12], that use Bayes theorem to relate the posterior probability $P(C|\mathbf{X})$ to a likelihood $P(\mathbf{X}|C)$, then assume independence of the features, X_i. In this approximation $P(\mathbf{X}|C) = \prod_{i=1}^{N} P(X_i|C)$ and so the contribution, $P(X_i|C)$, to the disease from each factor can be calculated observationally and studied separately.

3 Omes from an Ecological perspective

An ecological analogy may help intuit the difference between the two types of "ome". The traditional molecular "omes" are akin to an ecological community, where one is interested in the mutual interactions between all the structures in the system. In ecology these are typically species. The construction of a disease-ome, on the other hand, is more akin to the construction of an ecological niche, where, now, the disease itself is seen as a "species", as those factors that favour a high value for $P(C|\mathbf{X})$, relative, say, to a null hypothesis, $P(C)$, can be viewed as being niche-like, favouring the presence of the disease, C, while low values relative to the null hypothesis are anti-niche-like, favouring the absence of the disease. In this context, taking type 2 diabetes as an example, a niche factor may be the presence of a disease-related SNP, such as *rs*8050136 [13], as may be the consumption of carbonated drinks, or the price of carbonated drinks, or educational status, or hours of exercise, or age, or knowledge of the health consequences of diabetes or the health consequences of consumption of sugary foods, or a seemingly endless array of other factors. Unlike the human genome project there is, to our knowledge, no diabetes-ome project, where the goal is to obtain and integrate the multi-scale, multi-discipline data that begins to represent the totality of factors that affect the development of type 2 diabetes. Project 42, developed at the Centro de Ciencias de la Complejidad of the UNAM is a step in that direction in the context of obesity and metabolic disease. With over 3000 participants and several thousand variables, from a spectrum of previously identified SNPs for risk of obesity and metabolic disease, through demographic data, personal and family history, an ample set of biomarkers, anthropometric measurements, health knowledge, psychometrics, social characteristics, actigraphy and habits; all in a publicly available platform for analysis. The challenge of such data sets is to go beyond a static, statistical description to a process-oriented causal characterisation. Besides the right mathematical tools, this also requires domain-specific knowledge that spans multiple disciplines. The example of genetics affecting the impact of diet and diet affecting the expression of genes, while diet itself is a result of consumption and the consequence of a large number of other factors, from family environment to culture and mass marketing campaigns, speaks to the huge challenges of constructing a more process-oriented framework. Even just discovering the true underlying causal connections between, say, obesity and a single proxy variable such as educational level presents enormous

challenges. Thus, the construction of $P(C|\mathbf{X})$ via some suitable algorithm is just a first necessary step.

An advantage of a Bayesian framework for developing the disease-ome is that, based on a set of factors - "niche" dimensions - $\mathbf{X}$, it can be naturally extended by incorporating new information, such as new variables $\mathbf{X}'$. In this case, the posterior probability, $P(C|\mathbf{X})$, relative to the prior probability, $P(C)$, in the absence of the information $\mathbf{X}$ can, in its turn, be taken as a new prior probability and a new posterior probability, $P(C|\mathbf{X}'\mathbf{X})$, that incorporates information from both $\mathbf{X}$ and $\mathbf{X}'$ constructed using Bayes theorem: $P(C|\mathbf{X}'\mathbf{X}) = P(\mathbf{X}'|C\mathbf{X})P(C|\mathbf{X})/P(\mathbf{X}'|\mathbf{X})$. Again, $P(\mathbf{X}'|C\mathbf{X})$ can then be estimated using one of several Machine Learning methodologies. Dynamics can be incorporated by considering $\mathbf{X} \equiv \mathbf{X}(t)$, where the state vector $\mathbf{X}(t)$ may also contain historical information. For instance, $\mathbf{X}(t)$ may contain information about someone's historical activity level such as: actual activity level, activity level one year ago, activity level two years ago etc. [14]. Prediction in time follows naturally from considering the estimation of $P(C(t)|\mathbf{X}(t'))$, where $t' < t$. Thus, we may predict the probability for a disease state to occur at time t given the disease niche at t'. An example would be predicting if someone would become diabetic in a certain year given their disease-ome in previous years.

4 The Conductome

A process-oriented perspective requires us to think in terms of temporal development - of change. In the case of living systems, change is most naturally thought of in terms of *behaviour*. It is behaviour that naturally links cause and effect, with behaviour being the natural response - effect - to external or internal stimuli - causes. Indeed, one may argue that it is the sole medium by which organisms interact with their environment, including with other organisms. It is clear that our genome codes not only the physical structures of an organism but also what they do, up to the collective behaviour of the organism as a whole, and beyond, to the collectivity of groups of individuals. However, it should be equally clear that behaviour leaves an imprint on the genome. To a large degree, the survival of an organism in its environment is associated with what it does. A behaviour that is apt for a given environment will be propagated, genetically, if it has at least a partial genetic origin, or culturally.

The majority of, if not all, biologically relevant behaviours, such as sleeping, eating, reproducing, evading predators, are linked to clocks and underlying physical rhythms, such as the circadian rhythm, and are biological responses to fundamental properties of the earth's motion in space and time. These rhythms are a fundamental part of the niche of almost all living organisms. Thus, we would argue, that it is behaviour that is the most natural link between the micro-omes, such as the genome, and more macro-omes, that are proxied by those variables that are the area of interest of disciplines such as epidemiology, sociology and psychology. Behaviour is both caused by genetic structures and functions and, in turn, leads to changes in those structures and functions. In the omic spirit we have posited the "Conduc-

tome"[1] - the totality of behaviours of an organism and the causative factors associated with them - as the most comprehensive link between micro-omes, such as the genome, and macro-omes. As with the micro-omes, the Conductome can be approached using Complex Networks by, for instance, considering the "interactions" between different behaviours, as well as their links to other factors. Similarly, they may be considered, in analogy with a disease-ome, in terms of a probability function, $P(C|\mathbf{X})$, where C is the behaviour of interest and $\mathbf{X}$ the set of corresponding factors linked to or predictive of C.

Like many deep concepts, behaviour is difficult to characterise precisely. Take as an example, thermoregulation in mammals [15]. Mammals have different responses to external temperature. A human may sweat or may take off a coat. An elephant may flap its ears while a dog may pant. Should we classify ear flapping, panting and coat removal as behavioural adaptations and sweating as a physiological adaptation? Sweating and panting are both controlled by the autonomic nervous system [16, 17]. What about taking off a coat? We would argue that a more general and appropriate characterisation of behaviour, such as "the internally coordinated responses (actions or inactions) of whole living organisms (individuals or groups) to internal and/or external stimuli" [18] would naturally classify them all as behaviours. What about eating? Is eating more like sweating or taking off a coat? Both, of course. There is a basal mechanism for generating the urge to eat that comes from the autonomic nervous system [19]. This is a natural view from a biological perspective. However, there is much debate about just how "automatic" eating is [20], that is linked to polemical issues such as "fat shaming" and free will [21]. Clearly, as a fundamental necessity of life, the desire to seek and obtain food is pre-programed genetically. However, what about: what we eat? how much we eat? when we eat? where we eat? etc. When we eat, for instance, is associated now with an entire field of study - chrononutrition [22], while there is ample evidence that certain food preferences, such as fatty and sugary foods [23] have a neurobiological link [24]. How much we eat - portion size - is another dimension that has a strong psychological and social component, if not a direct biological one.

Although there exist ontologies of behaviour, such as the Neuro Behaviour Ontology [25], food consumption offers a good example of the complexity of classifying behaviour. We may consider food consumption as a behaviour as a class to be predicted. Obviously the simple class $C = food\ consumption = YES/NO$ is not useful, as all humans must consume food. Indeed there are a set of underlying basal behaviours that are intimately associated with the fundamental properties of life, such as homeostasis, metabolism, reproduction and adaptation to the environment. Food consumption is vital for metabolic processes and to maintain homeostasis. However, beyond the pure classification of consumption = YES/NO, we may construct a multitude of classes of interest following the discussion above. For instance, C may represent overconsumption, as defined with respect to some baseline null hypothesis, consumption of a certain food type, consumption as classified through

[1] The Conductome was introduced in the international workshop "The Human Conductome: A New Paradigm for Understanding Obesity? in the C3 ? Centro de Ciencias de la Complejidad, UNAM 29-30th November 2018.

portion size, consumption by eating times, consumption by frequency, consumption by place or, indeed, any and all combinations of the above, and more. By considering different classes we may determine the degree of heterogeneity associated with these classes/behaviours. The set of predictors, **X** - genetic, epigenetic, physiological, psychological, social - then represent the Conductome for that behaviour. Project 42, alluded to above, is, in this sense, an attempt to construct a set of variables across multiple scales that may begin to approximate in certain dimensions a Conductome for those behaviours - overconsumption and sedentariness - that are particularly related to obesity and metabolic disease. Additionally, if we do not have direct observations of a particular behaviour, we may imagine constructing a Conductome indirectly, by taking as a class C a physiological state, such as obesity or hypertriglyceridemia, that we hypothesise that is correlated with a behaviour of interest, such as overconsumption of food. Of course, when speaking of behaviour, there is a natural structural element - the central nervous system and its description at multiple spatial scales, from the cellular to the cortical - that must be included as an intermediary between the causes of behaviour and the consequences of behaviour. In this sense a "neurome" will be an essential element in understanding the link between cause and consequence and back again.

As a final reflection, although the workshop focused on physiological rhythms - periodic and periodic-like processes in human physiology - many life processes and corresponding behaviours are not periodic. Such periodicity is particularly relevant when we consider those short time intervals that are dominated by the earth's principal spatio-temporal rhythms: day, month or season, and is naturally linked to reversibility, if one thinks of returning to an initial state. However, life and its associated behaviours have a cost. Although one may analyse such costs in much more sophisticated terms - of information, entropy and free energy [26] - for present purposes one can think of the costs in terms of "wear and tear". Life inevitably leads to wear and tear, and this can be measured along two principle dimensions - temporal extent and rate. All else being equal, a human of 50 years of age will exhibit more wear and tear than a human of 20 years of age [27, 28]. Similarly, twenty years of chronic stress and inflammation due to morbid obesity will be associated with a much higher rate of wear and tear than twenty years of abstemious living. Wear and tear at the physiological level is a result of behaviour. Organisms must feed, organisms must reproduce, organisms must survive in an uncertain environment.

Acknowledgements This work has been supported by DGAPA PAPIIT grants IG101520 and IV100520, CONACyT Fronteras grant FC-2015-2-1093, SECITI grant 093/2018 and a donation from Microsoft Academic Relations. The author is grateful for very fruitful discussions with Peter Gollwitzer, Rolando Diaz-Loving, Ruud Buijs, Carolina Escobar, Sachin Panda, Per Sodersten, Cecilia Bergh, Lucia Ledesma and other participants at the first Human Conductome Workshop and to participants at the Matrix workshop Mathematics of Physiological Rhythms.

References

1. Ballereau S. et al. (2013) “Functional Genomics, Proteomics, Metabolomics and Bioinformatics for Systems Biology”. In: Prokop A., Csukás B. (eds) Systems Biology. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6803-1_1
2. Horgan, R.P. and Kenny, L.C. (2011) ?Omic? technologies: genomics, transcriptomics, proteomics and metabolomics. The Obstetrician & Gynaecologist, 13: 189-195. doi:10.1576/toag.13.3.189.27672
3. Haken, H. (2004) Synergetics. Introduction and Advanced Topics, Springer, Berlin.
4. Elena Kuzmin *et al* “Systematic analysis of complex genetic interactions” (2018) Science, Vol. 360, Issue 6386, eaao1729 DOI: 10.1126/science.aao1729.
5. Ivanov, P.C. and Bartsch, R.P. (2014) “Network physiology: mapping interactions between networks of physiologic networks.” In Networks of Networks: the last Frontier of Complexity (pp. 203-222). Springer.
6. Ivanov, Plamen Ch, Kang KL Liu, Aijing Lin, and Ronny P. Bartsch. (2017) “Network physiology: From neural plasticity to organ network interactions.” In Emergent Complexity from Nonlinearity, in Physics, Engineering and the Life Sciences, pp. 145-165. Springer
7. Ivanov, P., Zhang, X., & Lombardi, F. (2019) APS March Meeting Abstracts.
8. Lin, A., Liu, K.K.L., Bartsch, R.P. et al. Dynamic network interactions among distinct brain rhythms as a hallmark of physiologic state and function. Commun Biol 3, 197 (2020).
9. Black N, D’Souza A, Wang Y, et al. (2019) “Circadian rhythm of cardiac electrophysiology, arrhythmogenesis, and the underlying mechanisms”. Heart Rhythm; 16(2): 298-307. doi:10.1016/j.hrthm.2018.08.026
10. Erdmann, Jeanette et al. (2010) “Genetic causes of myocardial infarction: new insights from genome-wide association studies.” Deutsches Arzteblatt international vol. 107, 40: 694-9. doi:10.3238/arztebl.2010.0694
11. Zhe Wang, Bing Yu (2019) “Chapter 15 - Metabolomics, Proteomics, and Genomics: An Introduction to a Clinician”, Editor(s): Vijay Nambi, Biomarkers in Cardiovascular Disease, Elsevier, Pages 159-170, ISBN 9780323548359, https://doi.org/10.1016/B978-0-323-54835-9.00015-6.
12. Christopher R. Stephens, Hugo Flores Huerta and Ana Ruíz Linares (2018) “When is the Naive Bayes approximation not so naive?” Mach. Learn. 107:397?441 https://doi.org/10.1007/s10994-017-5658-0
13. van Hoek M, Dehghan A, Witteman JC, et al. (2008) “Predicting type 2 diabetes based on polymorphisms from genome-wide association studies: a population-based study”. Diabetes; 57(11): 3122-3128. doi:10.2337/db08-0425
14. Stephens C.R., Gutiérrez J.A.B., Flores H. (2020) Bayesian Classification of Personal Histories - An application to the Obesity Epidemic. In: Hassanien A., Azar A., Gaber T., Bhatnagar R., F. Tolba M. (eds) The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2019). AMLTA 2019. Advances in Intelligent Systems and Computing, vol 921.
15. Jeremy Terrien, Martine Perret, Fabienne Aujard (2011) “Behavioral thermoregulation in mammals: a review”, Frontiers in Bioscience 16, 1428-1444.
16. Nisha Charkoudian, “Chapter 60 - Hypothermia and Hyperthermia” (2012) Editor(s): David Robertson, Italo Biaggioni, Geoffrey Burnstock, Phillip A. Low, Julian F.R. Paton, Primer on the Autonomic Nervous System (Third Edition), Academic Press, Pages 287-289.
17. Krönert H, Pleschka K (1976) “Lingual blood flow and its hypothalamic control in the dog during panting”. Pflugers Arch. 1976; 367(1):25-31. doi:10.1007/BF00583652
18. Levitis, Daniel; William Z. Lidicker, Jr; Glenn Freund (2009) “Behavioural biologists do not agree on what constitutes behaviour”. Animal Behaviour. 78 (1): 103-10. doi:10.1016/j.anbehav.2009.03.018
19. Guarino D, Nannipieri M, Iervasi G, Taddei S and Bruno RM (2017) “The Role of the Autonomic Nervous System in the Pathophysiology of Obesity”. Front. Physiol. 8:665. doi: 10.3389/fphys.2017.00665

20. Adela R. Moldovan, Daniel David (2012) "Features of automaticity in eating behavior", Eating Behaviors, Volume 13, Issue 1, Pages 46-48.
21. David A Levitsky and Carly R Pacanowski (2012) "Free will and the obesity epidemic", Public Health Nutrition: 15(1), 126-141 doi:10.1017/S1368980011002187
22. Johnston, J. D., Ordovás, J. M., Scheer, F. A., & Turek, F. W. (2016) "Circadian Rhythms, Metabolism, and Chrononutrition in Rodents and Humans". Advances in nutrition (Bethesda, Md.), 7(2), 399-406. https://doi.org/10.3945/an.115.010777.
23. Adam Drewnowski, M.R.C. Greenwood (1983) "Cream and sugar: Human preferences for high-fat foods", Physiology & Behavior, Volume 30, Issue 4, Pages 629-633.
24. Allen S. Levine, Catherine M. Kotz, Blake A. Gosnell (2003) "Sugars and Fats: The Neurobiology of Preference", The Journal of Nutrition, Volume 133, Issue 3, Pages 831S?834S.
25. Musen MA, Noy NF, Shah NH, Whetzel PL, Chute CG, Story MA, Smith B; NCBO team (2012) "The National Center for Biomedical Ontology". J Am Med Inform Assoc. Mar-Apr;19(2):190-5. Epub 2011 Nov 10.
26. Erwin Schrödinger (1944) "What Is Life? : The Physical Aspect of the Living Cell". Based on lectures delivered under the auspices of the Dublin Institute for Advanced Studies at Trinity College, Dublin, in February 1943.
27. Stephens Christopher R., Easton Jonathan F., Robles-Cabrera Adriana, Fossion Ruben, de la Cruz Lizbeth, Martínez-Tapia Ricardo, Barajas-Martínez Antonio, Hernández-Chávez Alejandro, López-Rivera Juan Antonio, Rivera Ana Leonor (2020) "The Impact of Education and Age on Metabolic Disorders", Frontiers in Public Health 8, page 180; URL=https://www.frontiersin.org/article/10.3389/fpubh.2020.00180.
28. Antonio Barajas Martinez, Jonathan F. Easton, Ana Leonor Rivera, Ricardo Jesus Martinez Tapia, Lizbeth De la Cruz, Adriana Robles Cabrera, Christopher R. Stephens (2020) "Metabolic Physiological Networks: The Impact of Age", medRxiv 2020.08.05.20168997; doi: https://doi.org/10.1101/2020.08.05.20168997