시스템 생물학: 물리학과 생물학의 대화

DOI: 10.3938/PhiT.23.002

김 판 준

Systems Biology: Communication between Physics and Biology

Pan-Jun KIM

Driven by recent technologies to generate and store enormous amount of biological data, systems biology aims at achieving an integrative and quantitative understanding of biological systems from a holistic viewpoint. It provides both opportunities and challenges for physicists who want to tackle biological problems. In systems biology, physicists have to cope with the huge complexity of biological systems and need to be familiar with the concepts of evolution and phenotypes. Communication with biologists is highly recommended to resolve the challenges. On the other hand, as biological data continue to accumulate, systems biology will solicit more mechanistic insights that can be gained from physicists' analyses.

생명체? 과학자가 학문적 탐구 대상으로서 생명체라는 말을 들었을 때 떠올리는 생각 중 하나는, “복잡”하다라는 것이다. 생명체는 복잡하다. 생명체의 기본 단위로 보통 여겨지는 세포 하나도, 인간이 만든 그 어떤 정교한 기계보다 훨씬 더 복잡하다. 그리고 생명 현상은 다양한 시공간적 규모에 걸쳐서 일어난다(그림 1). 작게는 분자 수준에서부터 크게는 지질학적 사건에 이르기까지.

시스템 생물학(systems biology)은, 복잡한 생명 현상을 제

대로 이해하기 위해서는 생명체를 이루는 각 구성성분을 따로 떼어서 분석하는 것만으로 충분치 않다고 여기고, 구성성분들이 함께 어울려서 통째로 빚어내는 현상을 시스템 전체 수준에서 규명하는 것을 목표로 삼는다. 편의상 이 글에서는 현재 시스템 생물학의 주류 연구 대상이라 할 수 있는 세포 혹은 그 이하 대상을 염두하고 이야기를 진행하고, 필요에 따라 다른 대상들까지 이야기에 포함시키도록 하겠다.

각 세포 안에는 DNA로 된 유전자가 있고, 유전자의 염기 서열은 RNA를 거쳐 단백질 합성으로 이어져 단백질이 생체 활동을 수행하거나, 혹은 만들어진 RNA가 단백질 합성으로 이어지지 않고 직접 생체 활동을 수행하기도 한다. 유전자는 RNA와 단백질 합성의 설계도로서 생명 현상의 근간을 이루는 중요한 물질이고, 이 유전자라는 설계도의 내용을 가리켜 유전형(genotype)이라고 부른다. 한편, 직접 관찰이 가능한 특정한 생명 현상 자체를 가리켜서는 표현형(phenotype)이라고 한다. 표현형을 빚어내는 생체 분자들은 보통 혼자서 일하지 않고 함께 일하므로, 어떤 유전형에서 어떤 표현형이 나올지를 생체 분자 사이의 상호작용 및 환경의 영향을 고려해서 예측하는 것이 시스템 생물학의 일차 목표가 된다. 시스템 생물학의 최근 발전은, 대규모의 데이터를 얻는 생물학 실험 기술과, 이를 처리하고 공유할 정보통신 기술의 빠른 발달에 힘입은 바 크다. 과학자들은 실험으로 밝혀진 세포의 전체 유전자 염기 서열, RNA와 단백질의 발현량, 그 외 생체 분자들의 농도, 유전자 조절 네트워크, 단백질 상호작용 네트워크, 대사 네트워크, 신호전달 네트워크, 유전자 변이와 화학물질 처리에 따른 표현형의 변화 등을 인터넷을 통해 열람하고 공개할 수 있다. 이러한 데이터를 한데 엮어서, 유전형이라는 미시적인 성질에서 표현형이라는 거시적인 성질까지 원리 중심의 이해 체계를 세우는 데에 물리학자의 역할이 있다고 필자는 생각한다. 특히, 수많은 입자가 존재할 때 미시 세계의 동역학과 거시 세계의 관찰 결과를 확률, 통계적인 방법을 사용해서 연결하는 통계물리학은 시스템 생물학과 유사한 면이 있다. 그리고 시스템 생물학은 다루는 대상의 복잡성 때문에 정량 과학, 계산 과학이 될

저자약력

김판준 교수는 KAIST 물리학과에서 2008년 박사학위를 취득하였고, University of Illinois at Urbana-Champaign에서 박사후연구원(2008-2011)을 거쳐 현재 아시아태평양 이론물리센터에서 국제공동연구그룹 그룹장 및 POSTECH 물리학과 겸직교수로 재직하고 있다.
(pjkim@apctp.org)

Fig. 1. Various time and length scales covered by biological systems.

Fig. 2. Perspective of systems biology.

수밖에 없고, 이는 정성적인 결과 기술에 만족하는 기존의 여러 생물학 분야와 차별화되어 물리학자에게 다가온다(그림 2).

넘어야 할 고개들

물리학자에게 시스템 생물학은 기회이기도 하지만 동시에 어려운 도전이기도 하다. 앞에서 통계물리학을 잠깐 언급했지만, 생명체는 전형적인 통계물리학의 대상과는 또 다르다. 통계물리학에서 다루는 입자들은 균일하거나, 아니면 불균일하더라도 종류가 많지 않거나, 아니면 불균일성 자체에 근사적으로라도 수학적 체계가 있는 경우가 대다수이다. 이와는 다르게, 생명 현상에 관여하는 생체 분자들은 굉장히 불균일하고, 상호작용하는 방식도 서로 대등하기보다 계층 구조를 흔하게 이룬다. 이 때문에, 생명체를 이루는 개별 생체 분자나 생체 분자의 소모임은 자신만의 이야기거리를 풍부하게 가지는 경우가 많고, 이것이 왜 여러 생물학자가 개별적인 유전자나 생체 분자의 기능 연구에 그토록 초점을 맞추는지를 부분적으로 설명한다. 시스템 생물학에 입문하는 물리학자는 각 생체 분자의 기능에 대해 전문 생물학자만큼의 지식을 가지고 있지 않으므로, 여러 종류의 생체 분자를 일단 “익명화”시켜서 분석한 후 거기에서 그냥 끝내는 한계점을 보이는 경우가 있다. 이런 익명화 작업은 기존의 생물학자들이 미처 생각지 못했던 새로운 관점의 제시로 이어지는 경우도 있지만, 본질적으로 불균일한 생명 현

상을 다루는 데에 있어서 깊은 단계로 나아가지 못하고 오히려 잘못된 해석과 결론으로 이끄는 수도 있다. 결국 아는 만큼 보인다고, 자신이 연구하는 대상에 대한 생물학 지식을 쌓고, 생물학자와 자주 대화함으로써 그들의 시각과 정보를 공유하는 노력이 필요하다. 물리학자가 생물학을 할 때 어떤 일이 벌어질는지, George Cowan은 다음과 같이 묘사하였다.^[1]

“물리 과학은 개념적인 우아함과 해석적인 단순 명료함으로 대변되는 분야로 거의 정의되고 있다. 그러므로 당신은 그런 일을 거의 싫어도 해야 하고 다른 것들은 피해 나가게 된다…… 일단 생물학과 만나게 되면 당신은 그런 우아함을 포기해야 하며 또한 단순 명료함도 포기해야 한다. 모든 것이 혼잡스럽게 된다. 그리고 이로부터 경제학이나 사회 문제들로 쉽게 확산해 나갈 수 있다. 당신이 생물학에 조금이라도 몸을 담그게 되면 아마 빠져 나오기보다는 수영하기 시작하는 편이 나올 것이다.”

그리고 생명체가 전형적인 물리학의 대상과 또 다른 점은 진화한다는 것이다. 생명체는 진화라는 과정을 통해 항상 변해왔고 앞으로도 그럴 것이다. DNA 이중나선 구조를 밝혀낸 Francis Crick은 원래 물리학 출신으로, 생물학을 시작할 때 익숙하기 어려웠던 점이 바로 진화였다고 고백한 바 있다. 진화는 생명체가 현재의 모습대로 있도록 설계한 이치를 담고 있고, 각각의 특별한 물리, 화학적 법칙을 뛰어넘어 보다 보편적인 관점에서 생명 현상의 작동 원리를 바라볼 수 있게 해준다. 그리고 현재의 생물학 지식조차 진화의 과정을 거쳐 먼 미래에서는 바뀔 수 있다는 상념에 빠져들면, 결국 남는 것은 진

REFERENCES

- [1] M. M. Waldrop, *Complexity: The Emerging Science at the Edge of Order and Chaos*, Translated into Korean by K. Kim and H. Park, *From Chaos to Artificial Life* (카오스에서 인공생명으로) (Pumyang, Seoul, 1995).

화 자체가 작용하는 원리가 아닐까라는 생각이 든다. 이는 시스템 생물학에도 예외가 아니어서, 다루는 대상이 복잡할수록 그것을 꿰뚫는 이치에 대하여 생각해야 마음이 든든하므로, 시스템 생물학에 관심 있는 물리학자는 의식적으로라도 각 생명 현상을 진화와 연결시켜 생각하는 훈련을 해볼 것을 권유한다. 물리학자 특유의 강한 세계관을 바탕으로 진화에 대해 서술한 Nigel Goldenfeld의 최근 글도 일독을 권한다.^[2]

마지막으로, 생명체가 전형적인 물리학의 대상과 다른 점은 표현형에 있다는 것을 지적하고 싶다. 시스템 생물학을 대하는 일련의 물리학자들이 의외로 간과하기 쉬운 개념이 바로 표현형이다. 필자와 대화를 나누던 어떤 생물학자가 꺼낸 이야기가, 일부의 시스템 생물학 논문들이 공허한 느낌을 주는 이유는 바로 표현형의 부재 때문이 아닐까, 어떤 표현형을 보고자 하는 열망이 없기 때문에 무언가 공허한 단계에서 논문의 내용이 끝나버린 것이 아닐까라는 것이었다. 필자는 그 말에 일리가 있다고 생각한다. 특히 물리학자는 표현형까지 가지 않더라도 그 이전 단계의 묘사적인 수준에서 일을 끝내버리는 경우가 있는데, 시스템 생물학이 현실에 대해 강력한 예측력을 지니려면 생명체의 특정한 표현형까지 언급을 해줄 필요가 있다. 시스템 생물학은 정량 과학, 계산 과학을 넘어, 현실에서 검증받고 체감되어야 하기 때문이다.

기존의 성과

그럼 이제부터는 기존의 시스템 생물학 연구가 성취한 예들을 직접 살펴보도록 하겠다. 사실 시스템 생물학이라는 분야가 너무 커서, 여기에서는 물리학자에게 비교적 친화적인 예들을 중심으로 필자 주관대로 이야기를 풀어가겠다.

생명체의 표현형을 크게 결정짓는 것은 개별 생체 분자라기보다 생체 분자 사이의 네트워크라는 사실을 현대적인 의미에서 주장한 선구자 중에 Stuart Kauffman이 있다. 그는 젓당 대사에 관여하는 유전자 몇몇 사이의 조절 기작이 세상에 알려지자 이 개념을 확장하여 수많은 유전자가 다른 유전자의 발현을 서로 조절하는 거대한 유전자 조절 네트워크의 존재를 상상하고 전산 모사하였다. 각 유전자의 발현 상태는 다른 유전자의 발현 상태에 영향을 끼치거나 받으면서 네트워크 전체의 발현 상태가 시간에 따라 변화한다. Kauffman은 네트워크 전체의 발현 상태가 보이는 패턴이 해당 세포의 표현형과 대응할 것으로 생각하였다. 그는 이 일을 시스템 생물학이라는 분야가 대두되기 이전에 하였고, 그가 상상한 거대한 유전자 조절 네트워크는 실험 데이터가 있지 않은 상태에서 단순히 무작위인 구조를 가정하여 만들어낸 것이었다. 그럼에도 불구하고, 그가 보여준 생물학과 물리학의 경계를 넘나드는 통찰력

은 음미해볼 필요가 있다.^[3] 유전자 조절 네트워크가 시스템 생물학의 영역 안으로 들어오게 된 데에는 실제로 성계의 개체 발생에 관여하는 유전자 조절 네트워크를 실험으로 차례차례 밝히며 그 총체적인 모습에 접근해간 Eric Davidson의 역할이 크다. 그의 연구 그룹은 유전자 조절 네트워크의 발현 상태와 표현형의 관계를 명쾌하게 설명해왔고,^[4] 최근에는 자신들이 밝힌 유전자 조절 네트워크를 전산 모사가 가능한 수학적인 모형으로 제시하기도 하였다.^[5]

그 자신이 통계물리학자인 Albert-László Barabási가 이끄는 그룹은 생물학뿐 아니라 다른 분야의 네트워크까지 실제 데이터를 사용해서 분석하고 숨어있는 보편성을 찾아내어 본격적인 의미에서 네트워크 연구를 선도하였다.^[6] 생물학에서는 대사 네트워크와 단백질 상호작용 네트워크의 거대 구조를 선구적으로 분석하여 수많은 후속 연구를 이끌어냈다.^[7,8] Trey Ideker, Mark Gerstein, Marc Vidal 등이 이끄는 그룹은 생명체의 거대한 상호작용 네트워크 분석으로 유명하고, Uri Alon 그룹은 네트워크 모티프라는 개념을 창시하여^[9] 생명 현상을 만드는 단위는 개별 생체 분자라기보다 생체 분자들의 집합이라는 사실을 공고화하는 데에 앞장섰다.

대사 네트워크는 환경과 표현형의 관계를 비교적 명확하게 정량적으로 보여준다는 점에서 시스템 생물학의 중요한 연구 대상이다. 세포 전체의 대사 네트워크 구조를 밝히고 대사 물질의 흐름을 전산 모사할 수 있게 된 데에는 대사 공학이라는 실용적인 목적을 가졌던 화학공학 출신 과학자들의 역할이 컸다. Bernhard Palsson 그룹은 세포 전체의 대사 네트워크 구조 규명 및 대사 물질 흐름의 전산 모사 방법을 체계적으로

REFERENCES

- [2] N. Goldenfeld and C. Woese, *Annu. Rev. Condens. Matter Phys.* **2**, 375 (2011).
- [3] S. A. Kauffman, *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*, Translated into Korean by H. Kook, *The Edge of Chaos* (혼돈의 가장자리) (Science Books, Seoul, 2002).
- [4] E. H. Davidson, *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Academic Press, San Diego, 2006).
- [5] I. S. Peter, E. Faure, and E. H. Davidson, *Proc. Natl. Acad. Sci. USA* **109**, 16434 (2012).
- [6] A. -L. Barabási, *Linked: The New Science of Networks*, Translated into Korean by B. Kahng and K. Kim, *Link* (링크) (Dong Asia, Seoul, 2002).
- [7] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. -L. Barabási, *Nature* **407**, 651 (2000).
- [8] H. Jeong, S. P. Mason, A. -L. Barabási, and Z. N. Oltvai, *Nature* **411**, 41 (2001).
- [9] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman & Hall, Boca Raton, 2007).

정립하였고,^[10] 그 덕분에 다양한 주제에 걸친 수많은 후속 연구가 뒤따랐다.^[11,12] 예컨대, 그들이 정립한 방법론에다가 세포 내 효소들이 뻘뻘이 들어찬 정도를 접목하여 세포의 특이한 먹이 섭취 패턴을 설명한 Alexei Vazquez의 연구가 있는데,^[13] 생물학적 진실에 얼마나 부합하는지 여부를 떠나 시스템 생물학이 생물학의 기존 난제에 어떻게 도전할 수 있는지를 잘 보여주는 사례라고 생각한다.

지금까지 유전자 조절 네트워크, 단백질 상호작용 네트워크, 대사 네트워크 등의 예를 들었는데, 시스템 생물학이 추구하는 세포 전체의 모사가 정말하게 가능하려면 서로 다른 네트워크를 잘 합쳐서 그것들의 동역학을 통합적으로 기술해야 할 것이다. 상대적으로 매우 적은 수의 유전자를 가지는 특정 박테리아들은 이런 작업을 더 용이하게 할 수 있고, Markus Covert가 이끄는 그룹은 마이코플라스마 균을 대상으로 이 일을 정말로 성사시켰다. 그들은 해당 박테리아에 대해 알려진 모든 생체 분자들과 네트워크들을 종합하여 세포 분열이 한번 끝날 때까지 일어나는 모든 생화학적 과정을 전산 모사하였다. 그 결과, 기존에 알려지지 않았던 새롭고 흥미로운 현상들을 전산 모사를 통하여 관찰할 수 있었다.^[14]

아직 주류 시스템 생물학에 직접 적용되지는 않았지만, 생명 현상을 정보 처리의 과정으로 보고 분석하는 문제에 힘을 쏟는 William Bialek 그룹의 연구 결과도 주목해서 볼 만하다. 그들은 실험에서 관찰된 유전자 발현량 등이 생명체에서 정보의 흐름을 최적화할 수 있는 패턴을 따른다고 제안하였다.^[15] 그리고 시스템 생물학이라고 하면 거대한 네트워크를 꼭 다뤄야 하는 것이 아니라, 물리학자가 선호하듯이 시스템 전체를 소수의 변수로 근사하여 취급할 수 있음을 Terence Hwa가 이끄는 그룹이 유전자 발현량과 세포 성장 속도의 관계를 규명하여 밝혀냈다.^[16] 하지만 이런 근사도 생물학에 대한 튼튼한 이해가 전제되어 있기에 가능했다는 사실을 명심해야 한다.

전망

앞에서 이미 이야기했듯이, 시스템 생물학의 진정한 구현을 위해서는 여러 종류의 생체 분자 네트워크를 통합하여 분석하는 작업이 필요하다. 앞에서 이야기한 Markus Covert 그룹의 예에서 보듯이, 세포 전체의 동역학과 기능의 모사를 위해서 서로 다른 네트워크를 통합하는 작업은 점점 확대되어 나아갈 것이다. 그리고 세포 하나를 넘어서는 여러 세포의 집단을 시스템 생물학의 관점에서 연구하려는 노력도 더욱 필요하다. 이는 인체 생리학, 발생학, 면역학, 종양학에의 응용은 물론, 메타 유전체학의 발달로 각광받는 미생물 집단의 생태학 연구에도 도움을 줄 것이다. 이런 관점에서 필자의 연구 그룹은, 인

체에 서식하는 천여 종의 미생물과 인체 사이의 생태계를 분석하여^[17] 인체의 건강 및 질병과 어떤 연관이 있을지를 탐구하고 있다. 마지막으로, 합성생물학 (synthetic biology)의 발전은 시스템 생물학의 발전에도 큰 영향을 미칠 것이다. 이제 합성생물학은 소수의 외부 유전자를 세포에 옮겨 새로운 기능을 이식하는 수준을 넘어서서, 유전체 전체를 아예 새로 합성하여 세포에 통째로 옮겨 심는 단계에 이르렀다.^[18] 합성생물학이 펼쳐 보일 혁신적인 기술들은 시스템 생물학 연구에도 광범위하게 적용될 것이다.

이 글을 작성할 때 필자는 인간 유전체 해독을 천 달러 정도에 가능케 해주는 Illumina사의 제품 발표 소식을 들었다. 천 달러짜리 유전체 시대의 도래는 시스템 생물학이 기술의 최전선에서 극심한 데이터의 홍수를 겪을 것임을 다시 한 번 보여준다. 데이터가 홍수를 이룰수록 이를 원리적으로 이해하려는 노력은 더욱 절실하게 느껴질 것이다. “구슬이 서 말이라도 꿰어야 보배”라고, 시스템 생물학은 물리학이 필요하다.

ACKNOWLEDGMENT

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF-2012R1A1A2008925).

REFERENCES

- [10] B. Ø. Palsson, *Systems Biology: Properties of Reconstructed Networks* (Cambridge University Press, Cambridge, 2006).
- [11] P. -J. Kim, D. -Y. Lee, T. Y. Kim, K. H. Lee, H. Jeong, S. Y. Lee and S. Park, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13638 (2007).
- [12] C. B. Milne, P. -J. Kim, J. A. Eddy, and N. D. Price, *Biotechnol. J.* **4**, 1653 (2009).
- [13] Q. K. Beg, A. Vazquez, J. Ernst, M. A. de Menezes, Z. Bar-Joseph, A. -L. Barabási and Z. N. Oltvai, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12663 (2007).
- [14] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival, N. Assad-Garcia, J. I. Glass and M. W. Covert, *Cell* **150**, 389 (2012).
- [15] G. Tkačik, C. G. Callan, Jr, and W. Bialek, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 12265 (2008).
- [16] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang and T. Hwa, *Science* **330**, 1099 (2010).
- [17] P. J. Turnbaugh, R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight, and J. I. Gordon, *Nature* **449**, 804 (2007).
- [18] D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R. -Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z. -Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison III, H. O. Smith and J. C. Venter, *Science* **329**, 52 (2010).