

E-CELL 프로젝트 세포의 컴퓨터 시뮬레이션

Tomita Masaru/게이오대학교 환경정보학부

Genome, Proteome, Transcriptome 해석 등으로 대량의 생물학적 데이터가 축적되고 있으며, 이것을 기반으로 세포의 전체 활동을 모델화하기 위한 다양한 노력이 진행중이다. 본 고에서는 세포 시뮬레이션을 위한 소프트웨어 E-CELL과 이를 이용한 몇 개의 세포모델을 소개하고자 한다.

■ 서론

1980년대부터 세포 대사에 대한 컴퓨터 시뮬레이션이 연구되어 왔지만, cell cycle이나 유전자 제어 network 등 세포 process의 일부에 특정화 된 것 이었다. 세포는 수천, 수백만 종류의 단백질로 이루어져 있으며 서로 상

호 작용하는 매우 복잡한 시스템을 구성하고 있다.

불과 4, 5년 전까지만 해도 세포 자체를 시뮬레이션하는 것은 거의 불가능하다고 생각했으나, 최근 게놈연구와 컴퓨터 하드웨어의 급속한 진보로 많은 연구가 가능해 졌다. 즉, proteome 해석으로 세포 내의 모든 단백질의 기능구명이 가능하며, 현재의 컴퓨터 능력을 이용하면 수만 종류의 물질간의 상호작용을 시뮬레이션 할 수 있다.

I. E-CELL 프로젝트

미국 TIGR(The Institute for Genomic Research)연구소에서 Mycoplasma균(*Mycoplasma genitalium*)의 전체 게놈 서열을 동정·발

유전자형	Mycoplasma 유전자	기타	합계
해당	9	0	9
유산발효	1	0	1
인지질합계	4	4	8
인산전이효소계	2	0	2
글리세롤 수용	1	0	1
RNA polymerase	6	2	8
아미노산대사	2	0	2
리포좀 L subunit	30	0	30
리포좀 S subunit	19	0	19
rRNA	2	0	2
tRNA	20	0	20
tRNA ligase	19	1	20
개시인자	4	0	4
신장인자	1	0	1
단백질 코드 유전자	98	7	105
RNA 코드 유전자	22	0	22
합계	120	7	127

표 1 Virtual 세포 유전자수

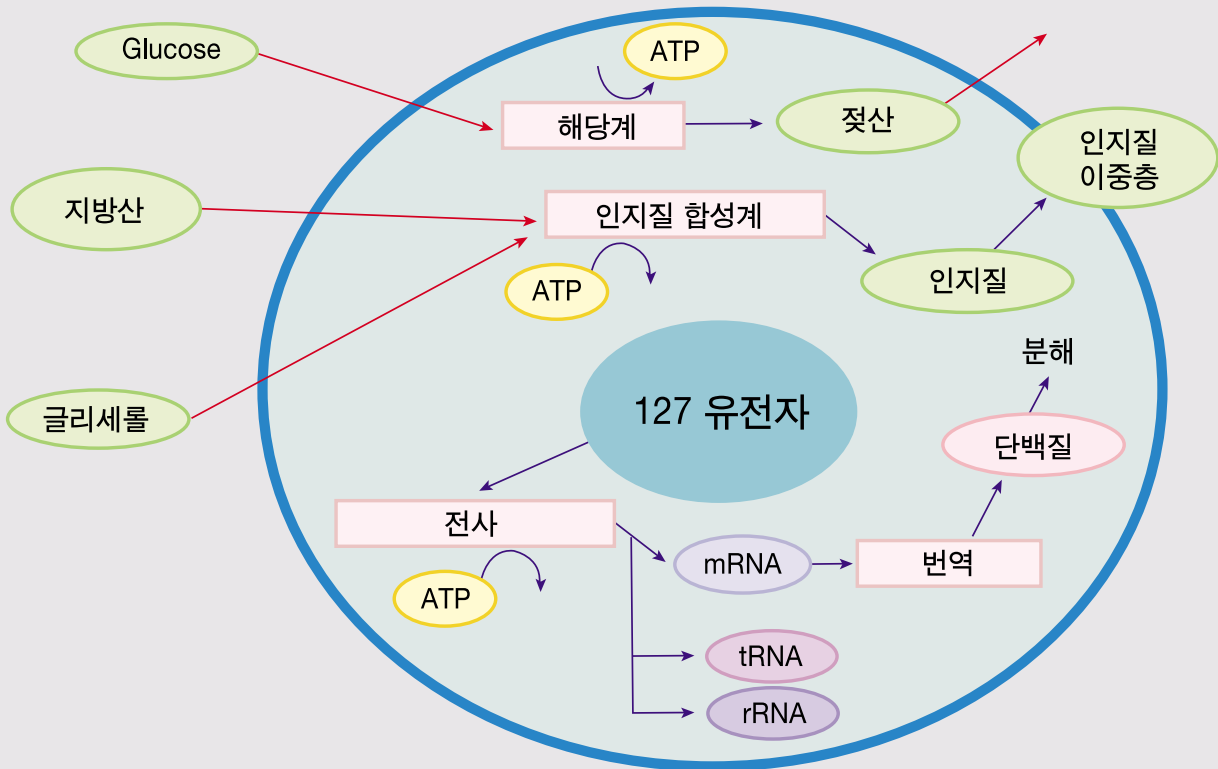


그림 1 E-CELL 시스템을 이용하여 구축한 "Virtual 자생 세포모델"

Virtual 자생세포는 127개의 유전자를 가지고 전사, 번역하여 단백질을 합성한다. 단백질 합성에 필요한 에너지(ATP)를 생산하기 위해서 막 외부에서 glucose를 섭취하여 해당계에서 분해한다. 또한, 세포막 생성을 위한 인지질 합성계를 가지며 지방산과 글리세롤을 이용하여 phosphatidylglycerol을 합성한 것이 세포막이다.

표한 직후, 1996년 게이오대학에서 세포 내의 대사 전체를 시뮬레이션하는 E-CELL 프로젝트를 수행하였다. *M.genitalium*은 현재 알려져 있는 생물 중에서 가장 게놈 길이가 짧고(58 kbp), 최소의 유전자 세트(약 480개)로 이루어져 있다. 이는 대장균의 약 10% 정도로, 세포 전체의 시뮬레이션을 위한 아주 적합한 모델 생물이다. 480개의 유전자 서열을 데이터베이스로 검색하면, 약 80%는 다른 생물과 상동성이 있어 유전자 기능을 예측할 수 있지만, 나머지 20% 정도는 그 기능을 알 수 없다. 그러나 20%의 기능을 알지 못한 채 전체를 재구축하기란 어려운 일이다. *M.genitalium*의 모든 유전자는 생존에 필수적인 것은 아니다. TIGR 연구소에서는 대부분의 유전자를 대상으로 knockout 실험을 한 후, 생존에 필수적인 유전자를 밝히고, 이 정보를 바탕으로 TIGR 연구소의 Clyde Hutchison, Craig Venter와 협력하여 세포 생존에 필수적인 최소한의 유전자 세트를 선별하여 가상의 "Virtual Cell"을 구축하고자 노력하고 있다.

먼저, E-CELL 프로젝트의 기반이 되는 시뮬레이션 소프트웨어 "E-CELL 시스템"을 개발하였다. 과거의 세포 시뮬레이션 연구에 이용했던 컴퓨터 프로그램은 각 과정에 특정한 소프트웨어로 호환성이 없었으나 E-CELL 시스템은 다양한 세포 프로세스(생합성계, 에너지 대사, 막 운

송, 전사, 번역, 복제, 시그널 전달 등)에 모두 대응할 수 있다. 따라서, 대사경로나 유전자 발현 제어 등을 동일한 형태로 모델화 할 수 있으며, 그 모델을 통합해서 세포전체를 시뮬레이션 할 수 있다.

1997년 여름, 미국 TIGR 연구소에서 세포가 생존하기 위해 필요한 최소한의 유전자가 127개임을 확인하였다(표 1). 그 중 *Mycoplasma* 게놈에 없었던 것은 7개로, 이는 대장균 등의 다른 종에서 유전자를 대체 한 것이다.

II. Virtual 자생 세포

그림 1은 E-CELL 시스템을 이용하여 구축한 "Virtual 자생 세포 모델"을 나타내고 있다. 이 Virtual 자생세포는 막 외부에서 glucose를 받아들여 해당계를 통해 에너지(ATP)를 생산한다. 또한, 세포막 생성에 필요한 인지질 합성계를 가지며, 지방산과 글리세롤을 이용하여 phosphatidylglycerol을 합성하는데, 이것이 세포막이 된다. 또한 유전자 발현을 위한 전사기구(RNA polymerase 등) 및 번역기구(ribosome 등)를 이용하여 유전자로부터 단백질을 합성한다. 단백질은 시간이 경과함에 따라 자연 분해되도록 모델화 되어 있으므로 단백질을 계속 만들지 않으면 세포는 죽고 만다. 이와 같이 단백질을 합성하기 위해서는 에너지(ATP)를 요구하므로 glucose가 필요하다.

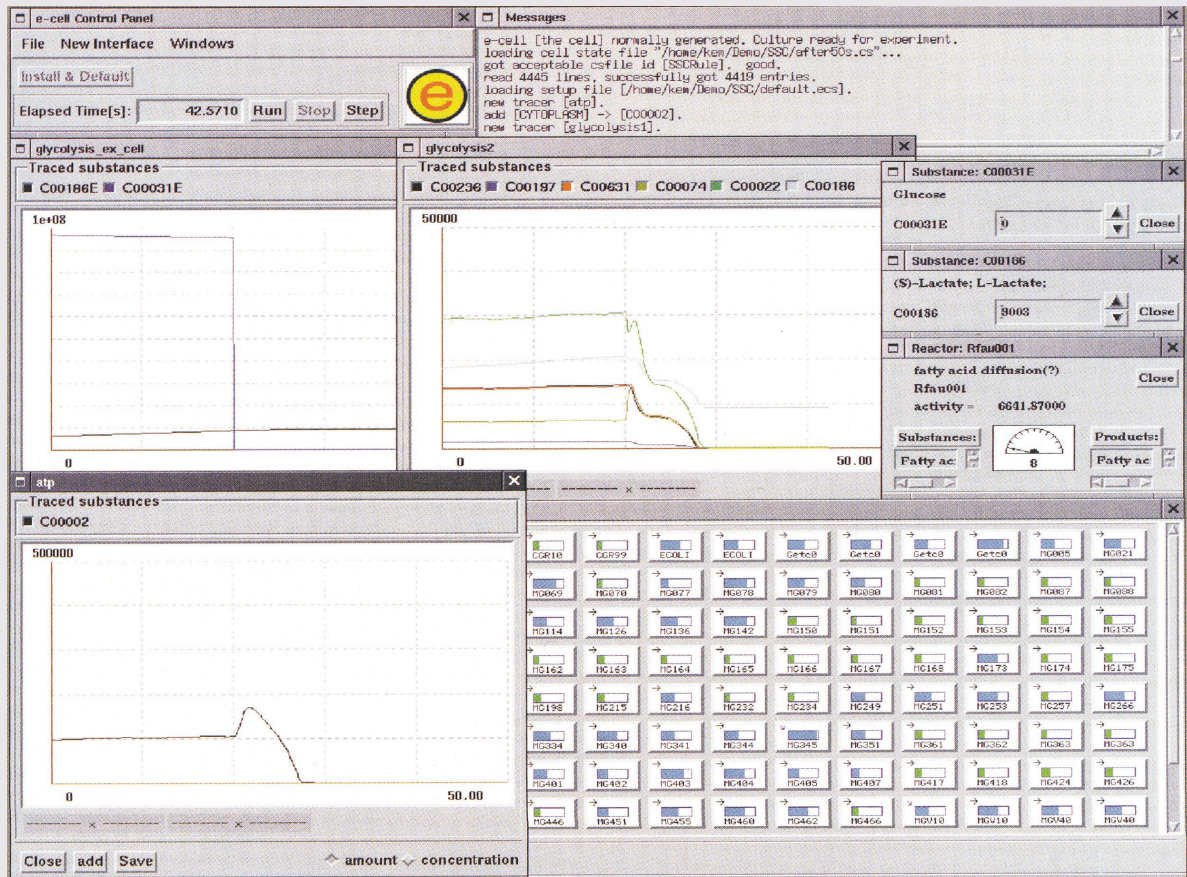


그림 2 E-CELL 시뮬레이션

세포내의 다양한 물질의 증감(분자수)을 관찰하거나("Traced substances") 특정 화학반응의 활성을 관찰할 수 있다("Reactor"), 모든 유전자의 발현상태가 일목요연하게 관찰되는 인터페이스도 있다(우측 하단). 또한 사용자가 실시간으로 물질의 양을 조절할 수 있다(우측 "Substance").

III. E-CELL 시뮬레이션 시스템

위에서 언급한 virtual 세포를 E-CELL 시스템을 이용하여 시뮬레이션 하면, 효소의 반응이 모두 병렬(유사병렬)로 실행되어 "대사활동"을 시작한다(그림 2). 그래픽 인터페이스를 통하여 세포 내 다양한 물질의 증감(분자수)(그림 2 "Traced substances")과 특정 화학반응의 활성도(그림 2 우측 "Reactor")를 관찰할 수 있다. 또한, 시뮬레이션 도중에 사용자가 개입하여 물질량을 조절할 수 있도록 되어 있다(그림 2 우측 상단 "Substance"). 모든 유전자의 발현 상태를 한눈에 볼 수 있다(그림 2 우측 하단).

각 아이콘이 각 유전자에 대응하고 있어 유전자의 전사량(mRNA의 분자수)을 나타내고, 마우스를 클릭하여 특정 유전자를 간단히 knockout 할 수 있다. Knockout 실험 중 세포의 움직임도 관찰할 수 있으며, 현실에서는 불가능한 실시간 knockout도 가능하다.

IV. Virtual 적혈구 세포

E-CELL 프로젝트로 구축한 자생 세포 모델은 가공의 세포로, 이를 근간으로 실제 세포를 시뮬레이션 해야 한다. 이 모델로 처음 선택한 것은 "인간 적혈구 세포"이다. 적혈구는 전사, 번역, 복제 등이 일어나지 않고 세포 내 대사가 제한되어 있어 가장 적합한 시뮬레이션 모델이며, 실험 데이터

가 풍부하여 컴퓨터 모델을 실제 세포와 비교 평가할 수 있다. 최근 E-CELL 시스템을 이용한 인간 적혈구 세포의 prototype이 완성되었고, 효소 기능을 의도적으로 저해하는 "Virtual 실험"을 통하여, 유전성 빈혈증 환자의 적혈구 상태를 재현할 수 있게 되었다. 그림 3은 인간 적혈구 E-CELL 모델이다.

현재 이 외에도 세포 소기관인 "미토콘드리아"의 모델화, 대장균의 주화성 (Chemotaxis: 유인물질의 농도가 높은 쪽으로 가는 성질)에 대한 시그널 전달계를 모델화 하고 있다.

V. 세포 시뮬레이션을 위한 metabolome 해석

E-CELL 프로젝트는 처음 시작할 당시 매우 미비하였으나, 최근에는 사이언스나 네이처를 비롯한 많은 잡지, 신문에 게재되는 등 많은 관심이 대두되고 있다. 구미를 중심으로 세포 시뮬레이션의 중요성이 급격히 인식되고 있으며, 미국에서는 에너지성과 국립 보건원(NIH)이 세포 모델화 프로젝트를 계획하고 있다.

향후 세포 시뮬레이션 연구는 점점 활성화 될 것으로 보이나, 정량적 데이터가 턱없이 부족하다(유전자의 기능을 알고 있다 하더라도 그들의 활동에 대한 정량적 데이터가 적다). 또한 세포 내의 다양한 물질의 농도에 대해서

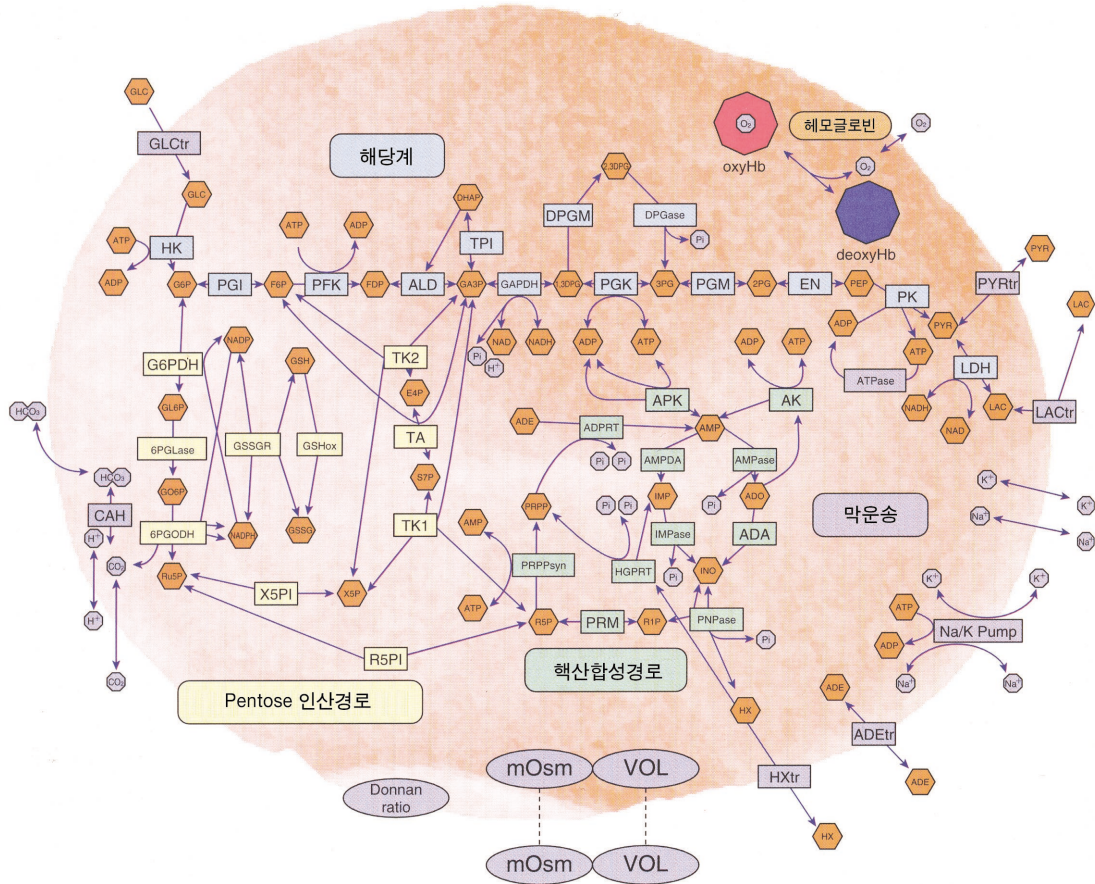


그림 3 인간 적혈구의 E-CELL 모델

적혈구 세포는 핵이 없으며 주로 3개의 경로 즉, 1) 해당계, 2) Pentose 인산 경로, 3)핵산 합성 경로가 있다. 이외에 효소 운반을 하는 헤모글로빈이 대량으로 존재하고 있다. 적혈구는 침투압 유지를 위한 대사를 담당하고, 헤모글로빈을 갖고 있는 주머니라고 할 수 있다.

도 전혀 알려진 바가 없으며, 적혈구처럼 세포 내 대사가 전체적으로 밝혀지는 경우는 극히 드물다.

지금부터는 metabolome, 즉 세포 내 대사의 정량적 데이터 해석이 매우 중요하며, 향후, “Metabolome+시뮬레이션”은 생명과학에 있어서 매우 중요한 역할을 할 것이다.

【참고문헌】

1) Tomita, et al: *Bioinformatics*(1999) 15: 72-84