

융합과학으로서 단백질 네트워크 Protein Network as a Fusion Science

김학용, 유재수 (충북대학교)

차 례

1. 서론
2. 단백질 네트워크의 응용
3. 단백질 네트워크의 진화
4. 단백질 네트워크와 신약개발
5. 결론

1. 서론

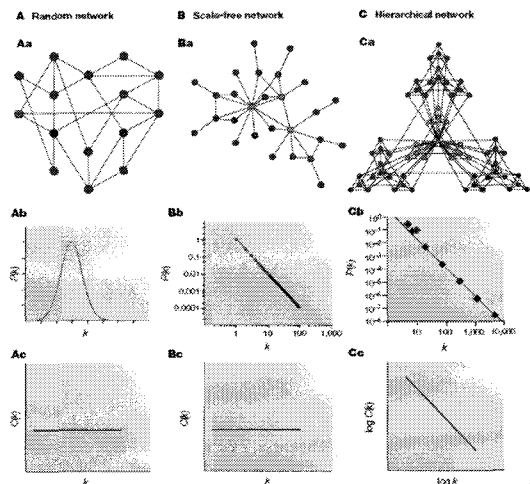
인간게놈프로젝트가 완성된 이후, 생명과학에서도 새로운 바람이 불기 시작했다. 하나의 유전자, 단백질, 대사물질을 다루는 고전적 접근 방법에서 총체적이라는 의미를 지닌 유전체, 단백질체, 대사체를 다루는 방향으로 급격히 선회하고 있다. 즉, 생명과학, 정보과학, 나노과학 등이 융합하여 생명현상을 규명하고, 인류에 기여하려는 학문적 기술적 목표를 달성하려는 시도이며 이를 융합과학이라고 한다.

융합과학의 일환인 네트워크 연구가 최근 들어 많은 관심을 불러 일으키면서 사회과학, 생명과학, 정보과학 분야에서 활발하게 응용되고 있다. 네트워크란 노드라는 점과 그들 사이를 연결하는 링크로 이루어진 하나의 기하학적 구조이다. 이러한 네트워크 연구는 솔로모노프와 라포포드[1] 및 에르되스와 레니[2]의 무작위 그래프 이론을 시작으로 사회학자 밀 그래프의 여섯 단계 분리[3], 와츠와 스트로가츠의 좁은 세상 네트워크[4]로 이어지고 있다.

최근에는 바라바시 등의 생명체 대사산물 네트워크[5]와 단백질 네트워크[6]로 이어지면서 생명과학 분야에서 네트워크 연구가 매우 활발하게 진행되고 있다. 생명체 네트워크에는 대사 네트워크[5], 단백질-DNA 네트워크[6], 단백질-단백질 네트워크[6], 유전자-질병 네트워크[7], 단백질-질병 네트워크[8], 약물-타겟 네트워크[9]에 관한 연구 결과가 발표되고 있다.

본 고에서는 주로 단백질-단백질 상호작용 네트워크를 중심으로 구조적 특성, 구조가 가지는 생물학적 의미, 네

트워크의 진화, 신약개발을 위한 도구로서 단백질 네트워크에 대해서 다루고자 한다. 생명체의 기원이나 질병의 근본적인 원인은 DNA라는 유전정보에서 출발하거나 찾는 것이 바람직하다. 그러나 질병의 타겟을 찾거나 네트워크로부터 유용한 정보를 얻기 위해서는 실제 생명체에서 생명활동을 주도하는 단백질을 대상으로 연구하는 것이 바람직하다.



▶▶ 그림 1. 네트워크의 분류

네트워크에는 크게 무작위 네트워크 (random network), 척도없는 네트워크 (scale free network), 계층적 네트워크 (hierarchical network)가 있다[그림 1]. 무작위 네트워크는 연결 수에 따른 단백질 누적 분포가 정규분포 (Poisson distribution)를 보여주고 있는데 반해[그림 1Ab], 척도없는 네트워크나 계층적 네트워크는

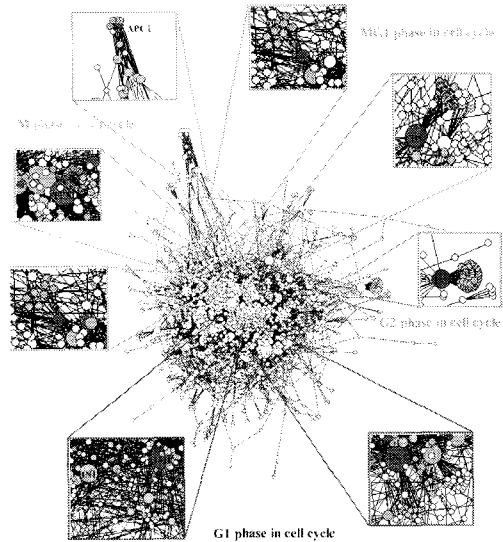
멱함수 분포(power-law distribution)를 보여주는 특성이 있다[그림 1Bb, Cb]. 척도없는 네트워크와 계층적 네트워크의 차이는 연결 수에 따른 단백질 누적 분포는 멱함수 분포로 같으나 연결 수에 따른 뭉침 계수의 분포에서 차이가 있다. 즉, 척도없는 네트워크에서는 변화가 없으나 계층적 네트워크에서는 연결 수에 따른 뭉침 계수 분포가 멱함수 분포를 가지는 것이 특징이다[10].

단백질 상호작용 네트워크는 척도없는 네트워크 및 계층적 네트워크의 특성을 동시에 가지고 있다[11]. 척도없는 네트워크는 연결 수가 매우 많은 비교적 소수의 허브 노드와 연결 수가 아주 적은 대부분의 노드로 구성되어 있기 때문에 무작위 네트워크에 비해, 정보의 전달이나 네트워크를 구성하는 노드들과의 상호작용이 매우 용이한 유리한 면이 있다. 그러나 이 경우 소수의 허브 노드가 중심이 되어 있기 때문에 허브가 외부 공격에 노출 시 네트워크의 와해가 일어난다는 약점을 안고 있다. 이에 반해 계층적 네트워크는 모듈구조를 가짐으로서 허브 중심의 중요한 정보를 모듈에 분산시킴으로서 외부 공격에 대한 네트워크의 견고성을 유지할 수 있는 특성을 지닌 구조를 가진다. 생명체뿐만 아니라 단백질 네트워크가 이러한 두 네트워크의 특성을 가지고 있는 것은 생명체가 환경과 세포 또는 세포와 세포사이에 정보전달을 빠르게 함과 동시에 외부의 공격에 노출될 수 있는 위험성으로부터 자신을 보존하여 생명체를 보존시키려는 특성을 동시에 반영할 수 있는 네트워크를 가진다고 할 수 있다.

2. 단백질 네트워크의 응용

2.1 단백질 기능적 모듈 형성

세포주기는 세포의 성장과 분열을 유도하는 매우 질서 정연하고 정교하게 진행되는 일련의 사건이다. 세포주기는 G1 → S → G2 → M기로 진행되는데, 세포주기의 각 기마다 특정한 단백질의 발현과 합성, 일련의 생화학적 인 경로를 통해서 이루어지고 있다[11]. 세포주기에 관여하는 단백질로부터 단백질 상호작용 네트워크를 구축하여 각 기마다 특징적으로 나타나는 모듈을 분석한다면, 유용한 정보를 획득할 수 있을 것이다. 그림 2는 효모의 세포주기 관련 단백질 상호작용 네트워크를 구축한 후 각 기에 특징적으로 나타나는 모듈을 분석한 것이다.



▶▶ 그림 2. 세포주기 관련 단백질 네트워크로부터 구축한 특징 기에 특이적으로 나타나는 모듈 분석.

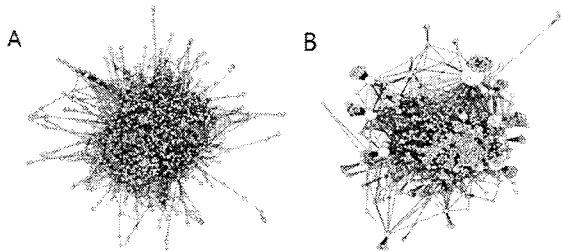
G1기에는 BIM1, BNI1, CDC45가 특이적 발현되어 이 기를 조절하고, G2기는 MEC3, M기는 APC1과 TEM1, Mrl와 MG1기에는 SPO12, MG1기에는 CHS1이 특이적으로 관여하는 것을 확인할 수 있다. 동시에 각 기에 특이적으로 관여하는 단백질 중심의 모듈을 형성하고 있음을 알 수 있다. 각 모듈은 모듈의 중심 단백질인 허브(붉은색 원), 이미 각 기에서 중요한 역할을 수행하는 것으로 알려진 단백질(노란색 원)과 각 기에서 어떤 역할을 하고 있는지에 대한 명확한 정보가 없는 단백질로 구성되어 있다. 각 모듈로부터 세포주기에 관여하는 단백질에 대한 추가적인 정보를 획득할 수 있을 뿐만 아니라 아직까지 기능이 밝혀지지 않은 단백질들의 기능을 예측할 수 있을 것이다.

2.2 단백질 기능 예측

네트워크를 통해서 얻을 수 있는 또 하나의 정보는 전산학적 접근 방법을 통해 아직까지 밝혀지지 않은 기능을 예측하는 일이다[12]. 가장 간단하고 일반적인 방법은 기능이 알려진 단백질과 상호작용하는 단백질은 같은 기능을 할 것이라는 가설을 근거로 예측하는 방법인데, 정확도가 낮은 것이 단점이다. 최근에 비교적 높은 정확도로 예측하는 방법이 발표되었다[13], 현재 단백질 및 상호작용 데이터베이스를 보여주는 있는 웹사이트에는 실제로는 상호작용하지 않는데 상호작용하는 것으로 나타나는(false positive) 데이터가 많이 존재한다. 진핵세포

는 여러 세포소기관을 보유하고 있다. 어떤 세포소기관에 존재하는 단백질은 다른 세포소기관에 있는 단백질과는 상호작용할 가능성이 희박하다. 따라서 서로 다른 소기관이 존재하는 두 단백질이 상호작용 데이터베이스에 있다면 제거함으로써 정확도를 높일 수 있다.

만약 같은 또는 유사한 기능을 하는 단백질들은 유사한 도메인(domain)을 가지고 있다는 가설을 단백질 기능 예측에 도입한다. 다시 말 해, 네트워크상에서 어떤 허브와 상호작용하는 단백질과 상호작용하는 단백질이 같은 기능을 하기 위해서는 최소한 하나 이상의 공통 도메인을 가지고 있을 것이다. 상호작용하는 단백질 모듈에서 최소한 하나 이상의 같은 도메인을 가지고 있지 않은 단백질을 제거 한 후 남은 단백질들 중에서 기능이 밝혀지지 않은 단백질의 기능을 부여한다면 정확도가 증가할 것이다[13].



▶▶ 그림 3. 효모 단백질 네트워크 A)정제하기 전의 단백질 네트워크 B) 세포소기관 및 도메인 정보를 근거로 정제한 단백질 네트워크

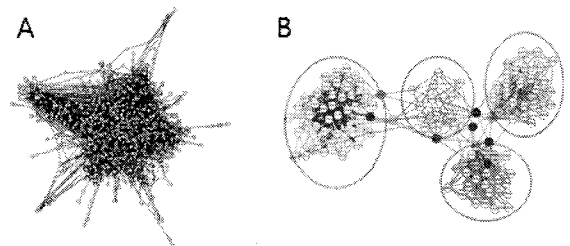
단백질 기능을 예측하는 알고리즘에서는 앞에서 언급한 단순히 이웃한 단백질에 기능을 부여하는 neighbor counting 방법이 있는데, 이는 이웃한 단백질의 거리에 상관없이 같은 가중치를 부여하고 뭉쳐져 있는 단백질의 크기를 고려하지 않고 있다. 또 다른 방법으로 Chi-square를 통한 단백질 기능을 예측하는 알고리즘을 들 수 있다[14]. 이 경우 기능 체계의 크기가 너무 크거나 작으면 효과적이지 못하며 또한 특정 단백질 중심의 모듈 크기에 영향을 받는다. 이 단점을 보완한 수정된 Chi-square 알고리즘에 의해 기능을 예측하는 알고리즘이 보고되었다[13]. 그림 3에서 보는바와 같이 세포소기관 정보와 단백질 도메인정보를 통해 데이터베이스를 정제한 효모 단백질 네트워크의 그림을 보여주고 있다. 이 네트워크를 바탕으로 수정한 Chi-square 알고리즘을 통해 80%의 정확도로 밝혀지지 않은 98개 단백질의 기능

을 부여하였다[13].

2.3 숨겨진 정보 발굴

복잡한 네트워크로부터 유용한 정보를 얻기는 쉬운 일이 아니다. 비록 형성된 네트워크를 구성하는 단백질에 기능, 소기관, 질병 연관성 등과 같은 정보를 주어서 기능 네트워크, 질병 네트워크, 소기관 네트워크를 구축하더라도 얻을 수 있는 정보에는 한계가 있다. 복잡한 네트워크(complex network)에서 가중치가 낮은 링크부터 차례로 제거하여 핵심 네트워크(core network)를 만든다면 좀 더 용이하게 네트워크를 이해하고 유용한 정보를 좀 더 쉽게 획득할 수 있을 것으로 생각한다.

가중치가 적은 노드를 우선적으로 복잡한 네트워크로부터 제거하는 방법의 일환으로 k-core 알고리즘을 들 수 있다[15]. k-core 알고리즘은 가중치가 낮은 링크를 반복적으로 제거함으로써 네트워크의 핵을 축으로 계층적 네트워크를 만드는데 매우 유용한 그래픽 도구라 할 수 있다. 단백질 네트워크를 포함한 생명체 네트워크는 척도없는 네트워크이면서 동시에 계층적 네트워크의 특성을 가지기 때문에 복잡한 네트워크를 핵심 네트워크로 전환하는데 이 알고리즘을 적용하는 것은 매우 적절하다고 생각한다.



▶▶ 그림 4. 효모 세포질 단백질 네트워크 A) 세포질에 있는 단백질 전체 네트워크 B) k-core 알고리즘을 사용하여 가중치가 낮은 링크를 제거한 후 구축한 core(핵심) 네트워크

네트워크에 숨겨진 정보를 찾기 위한 일환으로 효모 세포질에 있는 단백질 중심으로 단백질 네트워크를 구축하고 k-core 알고리즘을 사용하여 다시 핵심 네트워크를 구축하였다[그림 4]. 복잡한 네트워크에서 각 노드의 평균 65% 링크수가 제거되어 핵심 네트워크가 구축되었다. 네트워크에서 각 모듈을 연결시켜주는 노드를 병목 노드(bottle-neck node)라고 한다. 흥미롭게도 그림 4B

에서 각 모듈을 연결시켜 주는 3개의 병목 노드(GTS1, ARP2, BZZ1)는 약 9%~17%의 링크가 제거되었는데, 이는 다른 노드에 비해 대부분의 링크가 제거되었다고 볼 수 있다. 이들 단백질들은 실제 세포질 단백질 네트워크에서는 허브 단백질인데, 핵심 네트워크에서는 병목 노드로 재등장하고 있다. 이에 대한 해석을 위해서는 좀 더 네트워크의 진화과정과 메커니즘을 통해 살펴보아야 하겠지만, 외부의 공격에 취약한 허브 단백질을 보호하는 의도에서 실제로는 병목 노드인데, 덜 중요한 링크를 과도하게 연결시켜 허브로 위장함으로써 생명체의 견고성을 유지하려는 진화 산물일 가능성을 조심스럽게 제시한다.

3. 단백질 네트워크의 진화

초기 생명체에서 고등 동물로 진화하면서 단백질 수도 약 500개에서 사람인 경우 약 10만개로 증가하고 있다. 단백질 네트워크도 초기 생명체에 비해 사람의 네트워크는 더욱 복잡하다. 앞에서 언급한 바와 같이 이러한 복잡한 네트워크에서 유용한 정보를 찾는 것은 쉬운 일이 아닐 것이다. 이를 해결하는 일환으로 네트워크가 어떻게 생성되어 지금과 같이 복잡하게 진화되었는지에 대해 연구를 수행한다면, 이 과정을 통해서 생명 현상을 이해하고 나아가서 유용한 정보를 얻을 것이다.

단백질 네트워크의 진화를 설명하는 메커니즘으로 유전자 중복(gene duplication)을 들 수 있다. 생명체에서 유전자 중복은 안정한 DNA를 구축하기 위해 필연적으로 일어난다. 쌍을 이룬 유전자가 새롭게 합성되어 기존의 단백질 상호작용을 통해 단백질 네트워크를 고수하기도 하고, 때론 기존의 단백질 네트워크에서 벗어나 새로운 네트워크를 형성하기도 하며 기능도 다르게 바뀌기도 한다. 두 경우 진화를 계산하면 유전자 중복 후 기존의 네트워크를 유지하는 유전자보다 떨어져나가 새로운 네트워크를 형성하는 유전자의 진화 속도가 빠른 것으로 나타났대[16].

허브 단백질을 보유하고 있는 척도없는 네트워크가 형성되기 위해서는 새로운 단백질이 만들어질 때 어떤 단백질에 결합하느냐에 대한 가중치가 다르게 나타나야 할 것이다. 이를 근거로 제시한 모델이 선택적 결합(preferential attachment)이다[17]. 이는 이미 링크수가

많은 단백질에 우선적으로 결합하려는 경향성을 고려한 것으로 척도없는 네트워크와 부합되기도 하지만, 인터넷에서 먼저 탄생한 포털 사이트인 야후가 링크수가 많음에도 불구하고 후발 주자인 구글이 야후를 능가한 사실을 설명할 수 없는 단점이 있다. 생명체에서도 생명체의 다양성 및 모듈 구조를 가지는 계층적 네트워크를 설명하기는 미흡하다.

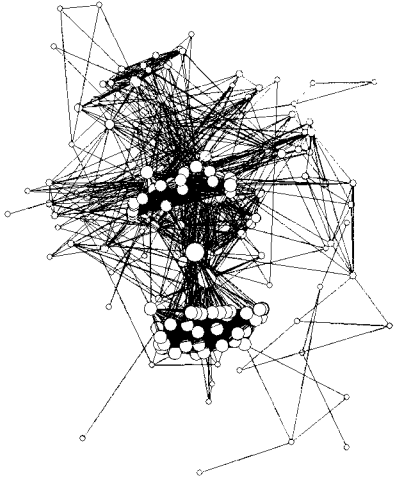
최근에 제안된 결정체 성장(crystal growth)모델이 있다[18]. 이 모델의 핵심 아이디어는 첫째, 상호작용하는 단백질 표면에 이용할 수 있는 공간의 존재 여부가 한 인자(factor)이다. 선택적 결합보다는 오히려 반 선택적 결합(anti-preferential attachment) 규칙을 고려한 것이다. 둘째, 네트워크가 성장함에 따라 연결 수가 많은 하위 네트워크(sub-network)가 단백질 모듈로 나타나 복잡계 네트워크를 만든다는 아이디어이다. 셋째, 일단 새로운 단백질이 한 모듈에 결합하면 그 다음 연결은 그 모듈을 중심으로 지역적으로 연결 수가 증가하는 경향성을 고려한 것이다. 이 모델을 시간의존성 단백질 네트워크의 진화라 일컫는다[18].

네트워크의 진화과정은 아직까지 초보단계이며 유전자 분석 위주로 진행되기 때문에 처음 소개한 유전자 중복 이론에 대한 연구가 중점적으로 진행되어온 것은 사실이다[16]. 본 고에서 제안하고자 하는 것은 이러한 연구를 필연적으로 진행하지만 동시에 네트워크 성장 모델 컴퓨터 프로그램을 개발하여 실제 단백질 네트워크와 비교 분석함으로써 실제 생명체에서 나타나는 진화과정을 시뮬레이션하는 방안도 고려할 필요가 있을 것으로 생각한다.

4. 단백질 네트워크와 신약개발

4.1 유전자와 질병 네트워크

인간 질병은 독립적으로 발생하거나 독자적으로 질병으로서 기능을 수행하기 보다는 유전자나 단백질과 마찬가지로 질병-질병 상호작용 네트워크를 형성하고 있을 가능성이 매우 높다. 질병 관련 유전자 정보는 OMIM(Online Mendelian Inheritance in Man)에서 확보할 수 있다[19]. 먼저 질병에 관여하는 유전자 네트워크를 구축 한다[그림 5].

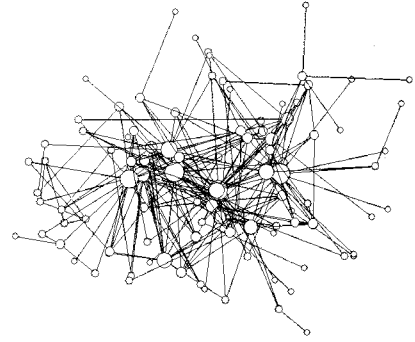


▶▶ 그림 5. 질병에 관여하는 유전자 상호작용 네트워크

유전자 네트워크에서 각 노드에 유전자 대신 질병명으로 치환한다. 만약 그 유전자가 질병에 관여하지 않는다면 그 노드는 질병 네트워크에서 사라질 것이며, 두 개 이상의 질병에 관여한다면 유전자 네트워크상에서는 하나의 노드가 질병 네트워크에서는 질병에 관여하는 질병 수만큼 증가할 것이다. 이러한 방법을 이용하여 질병-질병 상호작용 네트워크로 전환한다[7].

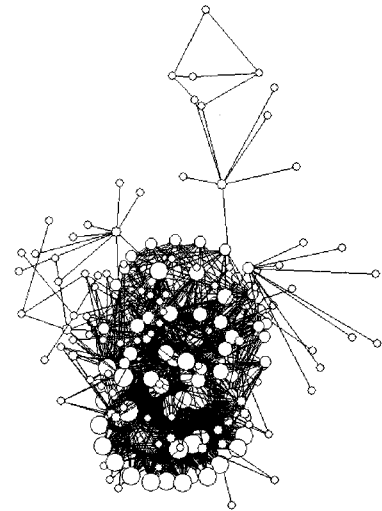
4.2 단백질과 질병 네트워크

유전자와 같은 방법으로 OMM에서 얻은 질병 관련 유전자를 먼저 단백질로 전환하고 이들 단백질로부터 단백질-단백질 상호작용 네트워크를 구축한다[그림 6]. 구축한 단백질 네트워크의 각 노드에 단백질 대신 질병명으로 치환하여 질병-질병 네트워크를 구축한다[8]. 유전자로부터 구축한 질병 네트워크와 단백질로부터 확보한 질병 네트워크에서 충분히 차이를 보일 수 있다. 왜냐하면 유전자의 경우, 질병에 관여하지만 실제 형질로 나타나지 않고 형질을 보존하는 유전자가 존재하는데 반해, 단백질의 경우는 질병 관련 유전자가 반드시 발현되어 나타나서 실제 질병에 관여하기 때문에 다를 수가 있다. 또한 유전자의 경우 양자 선택적 유전자 접합(alternative splicing)에 의해서 다양한 단백질을 만드는데, 그 정보가 아직 부족하다면 단백질 수준에서의 종류나 다양성에 의해서 차이점이 나타나기도 한다.



▶▶ 그림 6. 질병에 관여하는 단백질 상호작용 네트워크

4.3 약물 타겟 네트워크



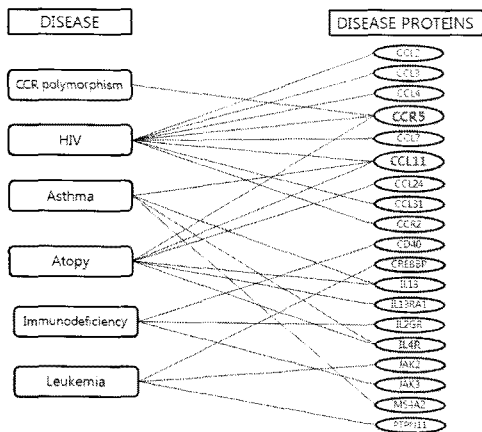
▶▶ 그림 7. 약물 타겟 네트워크

DrugBank database[20]로부터 질병치료에 이용되고 있는 약물의 실제 타겟이 되는 유전자 및 단백질 정보를 확보한다. 이들 정보를 바탕으로 질병 관련 유전자와 이 유전자를 타겟으로 하는 약물로 이루어진 네트워크를 약물 타겟 네트워크로 정의한다[9]. 암에 관여하는 유전자와 이 유전자와 상호작용하는 약물 타겟 네트워크를 구축하였다[그림 7]. 유전자 상호작용 네트워크나 단백질 상호작용 네트워크에는 각각 유전자와 단백질이 노드이며 유전자-유전자 또는 단백질-단백질 상호작용이 링크인데 반해, 약물 타겟 네트워크에서는 유전자 및 약물이 노드가 되고 유전자-약물 상호작용이 링크가 된다. 이러한 약물 타겟 네트워크로부터 현재 약물 데이터베이스에 구축된 타겟 유전자 또는 단백질을 확보할 수 있다. 따라서 유전자 기반 질병 네트워크, 단백질 기반 질병 네트워크 및 약물 타겟 네트워크에서 질병 단백질이나 질병

유전체를 확보할 수 있다. 약물 타깃, 질병 단백질, 및 질병 유전체를 상호 비교하여 효율적인 약물 타깃 단백질을 선정하고 선도화합물 도출을 위한 스크리닝 할 수 있는 가능성을 제시하고자 한다.

4.4 질병 단백질 발굴

약물 타깃 단백질 및 질병 단백체를 발굴하기 위해서는 질병과 관련된 단백질 또는 약물과 단백질 사이의 상관관계를 보여주는 세트가 요구 된다[그림 8]. 각 질병에 관여하는 단백질과의 연결고리를 보여주는 질병-단백질 이분할 그래프(bipartite graph)를 구축하여 대량으로 질병 단백체를 확보할 수 있다[7].



▶▶ 그림 8. 질병-단백질 이분할 그래프와 질병 단백질 확보

일반적으로 단백질 네트워크에서 허브 단백질들은 필수 단백질인 경향성이 높다. 질병-단백질 이분할 그래프로부터 얻은 질병 단백질들이 허브 단백질일 경향성이 있는 것으로 나타났다[8]. 단백질 네트워크에서의 허브는 네트워크의 견고성, 필수성, 질병과의 연관성과 같은 정보를 모두 포함하고 있기에 이에 대한 연구의 필요성이 대두된다.

5. 결론

21세기는 한 분야의 학문이 독주체제를 구축하는 그런 시대가 아닌 관련 학문들이 융합하여 새로운 트렌드를 추구하는 방향으로 진행할 것이다. 특히, 바이오기술(BT), 정보기술(IT), 나노기술(NT)이 어우러져 융합하는 학문 분야인 복잡계 과학(complex science)과 시스템생물학(systems biology)이 주도할 것이다.

이러한 융합과학의 하나인 단백질 네트워크에 대한 기본 지식에서부터, 네트워크의 응용성, 유용한 정보를 도출하는 방법, 네트워크의 진화에 이어 질병네트워크 및 질병 단백질 발굴에 이르는 분야를 여기서 다루었다.

지난 10년 동안 단백질 네트워크 분야에 관한 연구가 급격히 증가하고 있다. 2000년에는 일 년 동안 이 분야에 발표되는 논문수가 약 150여 편인 것이 2008년에는 약 700여 편이 넘는 논문이 발표되기도 하였다. 그러나 현재 일부에서는 네트워크 연구가 한계에 도달했다는 견해가 있는 것도 사실이다. 이는 복잡한 네트워크에서 노드는 단백질이고 링크는 방향성도 없는 단순 상호작용이기 때문에 더 이상의 유용한 정보의 도출은 한계에 도달했다는 시각에서 출발 한다.

노드인 단백질에 세포소기관 정보, 도메인 정보, 기능 정보, 발현 부위 정보 등을 첨부하여 다양한 각도로 분석을 하지만 더 이상의 유용한 정보를 얻기가 힘든 것처럼 보이는 것도 사실이다. 그러나 앞으로 단백질 네트워크 분야에서 가장 핵심적인 분야는 첫째, 단백질 네트워크의 진화 과정을 밝히는 것이다. 초기 생명체의 소수 단백질에서 진화를 통해 더욱 복잡하고 효율적인 방향으로 진화한 인간 단백질 네트워크를 놓고 그 진화과정을 밝히는 것이 쉬운 일이 아니지만, 생명의 근원 문제를 들여다보는 계기가 될 수 있을 지도 모른다. 둘째, 복잡한 네트워크를 형성하기 전 각 sub-network인 모듈들이 결합하면서 어느 시점에 창발적 행위(emergent behaviour)를 보여주는지를 밝히는 것도 매우 흥미 있는 일이다. 이를 위해서는 현재의 생물학, 정보 과학뿐만 아니라 프랙탈 기하학을 포함한 수학, 카오스 이론을 포함한 비선형 동역학과 같은 학문이 융합하여 그 실마리를 찾는다면 우리 앞에 놓인 이러한 근원적인 문제가 해결될 수 있을 것이다.

감사의 글

본 논문은 교육과학기술부와 한국산업기술재단의 지역혁신인력 양성 사업으로 수행된 결과임.

참고문헌

[1] R. Solomonoff and A. Rapoport, "Connectivity of random nets", Bull. Math. Biophys., Vol.13, No.1

- pp.107-117, 1951.
- [2] P. Erdős and A. Rényi, "On random graphs", *Publicat. Math.*, Vol.6, No.2 pp.290-297, 1959.
- [3] S. Milgram, "The small world problem", *Psych. Today*, Vol.1, No.1, pp.60-67, 1967.
- [4] D. Watts and S. Strogatz, "Collective dynamics of 'small-world' networks", *Nature*, Vol.393, No.6684, pp.409-410, 1998.
- [5] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabási, "The large-scale organization of metabolic networks", *Nature*, Vol.407, No.6804, pp.651-654, 2000.
- [6] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, "Lethality and centrality in protein networks", *Nature*, Vol.411, pp.41-42, 2001.
- [7] K.-I. Goh et al., "The human disease network", *Proc. Nat'l Acad. Sci.* Vol.104, pp.8685-8690, 2007.
- [8] 이윤경, 여명호, 강태호, 유재수, 김학용, "아토피관련 질병 네트워크로부터 질병단백체 발굴", *한국콘텐츠학회논문지* Vol.9, pp114-120, 2009.
- [9] M.A. Yildirim et al., "Drug-target network", *Nature Biotech.* Vol.25, pp1119-1126, 2007.
- [10] A.-L. Barabási and Z.N. Oltvai, "Network biology: understanding the cell's functional organization", *Nat. Rev. Genet.* Vol.5, pp101-113, 2004.
- [11] H.Y. Kim et al., "Phase specific activated modules from a protein interaction network of yeast cell cycle", *J. Kor. Phys. Soc.* Vol.50, pp290-294, 2007.
- [12] B. Schwikowski, P. Uetz, and S. Fields, "A network of protein-protein interaction in yeast", *Nat. Biotech.* Vol.18, pp1257-1261, 2000.
- [13] J.W. Ryu, H.Y. Kim, T.H. Kang, J.S. Yoo, and J.S. Chung, "Prediction of unannotated proteins from a protein interaction network filtered by using localization and domains in yeast", *J. Kor. Phys. Soc.* Vol.51, No.5, pp.1805-1811, 2007.
- [14] H. Hishigaki, K. Nakai, T. Ono, A. Tanigami, and T. Takagi, "Assessment of prediction accuracy of protein function from protein-protein interaction data", *Yeast*, Vol.18, pp.523-531, 2001.
- [15] I. Alvarez-Hamelin et al., "Large scale networks fingerprinting and visualization using the k-core decomposition", *Adv. Neu. Inform. Proc. Sys.* Vol.18, pp41-50, 2006.
- [16] T. Makino and T. Gojobori, "Evolution of protein-protein interaction network", *Gene. Prot. Evol.* Vol.3, pp13-29, 2007.
- [17] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks", *Science* Vol.286, pp509-512, 1999.

- [18] W.K. Kim and E.M. Marcotte, "Age-dependent evolution of the yeast protein interaction network suggests a limited role of gene duplication and divergence", *PLoS Comp. Biol.* Vol.4, ppe1000232, 2008.
- [19] OMIM <http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim>
- [20] DrugBank <http://www.drugbank.ca/>

저자소개

● 김 학 용 (Hak Yong Kim)

중신회원



- 1985년 2월 : 충북대학교 농화학(농학사)
- 1987년 2월 : 충북대학교 화학과 (이학석사)
- 1994년 5월 : 미국 코네티컷대학교, 분자 및 세포생물학과(이학박사)
- 1998년 ~ 현재 : 충북대학교 생화학과 교수

<관심분야> : 시스템생물학, 신호 전이, 단백질 네트워크, 생체동역학

● 유 재 수 (Jae-Soo Yoo)

중신회원



- 1989년 2월 : 전북대학교 컴퓨터공학과(공학사)
- 1991년 2월 : 한국과학기술연구원 전산학과 (공학석사)
- 1995년 2월 : 한국과학기술연구원 전산학과 (공학박사)
- 1996년 ~ 현재 : 충북대학교 정보통신공학과 교수

과 교수

<관심분야> : 데이터베이스 시스템, 정보검색, 멀티미디어 데이터베이스, 분산객체 컴퓨팅, 생물정보학