

류마티스관절염의 시스템 생물학적 접근

김기조¹ · 황대희² · 김완욱¹

가톨릭대학교 의과대학 성빈센트병원 류마티스내과¹, 대구경북 과학기술원 뉴바이올로지 전공²

Systems Approach to Rheumatoid Arthritis

Ki-Jo Kim¹, Daehee Hwang², Wan-Uk Kim¹

Division of Rheumatology, Department of Internal Medicine, St. Vincent's Hospital, The Catholic University of Korea¹, Suwon, Center for Systems Biology of Plant Senescence and Life History, Daegu Gyeongbuk Institute of Science & Technology², Daegu, Korea

Phenotypic characteristics of complex diseases such as rheumatoid arthritis are a consequence of interactions of genetic and environmental factors. Biomolecules closely interact with other molecular components and form functional modules, resulting in significant biologic action capability. While traditional biochemical research focuses on a single disease using narrowly constrained data, systems biology aims to interpret large volumes of highly complex and multilevel data obtained from high-throughput technologies to understand how biological systems function as a whole. Such a systems approach to complex diseases, so-called network

medicine, can shape our comprehensive understanding of disease mechanisms by identifying modules temporally and spatially perturbed in the context of health and diseases. Given the unmet needs for diagnosis, monitoring, and treatment in rheumatoid arthritis, systems biology is obviously an emerging powerful tool to gain insight into disease mechanisms, study comorbidities, analyze therapeutic drugs and their targets, and discover novel network-based biomarkers. **Key Words.** Systems biology, Network medicine, Rheumatoid arthritis

복합질환의 동질성과 이질성

유전질환(genetic disease)은 일반적으로 멘델질환(Mendelian disease)과 복합질환(complex disease)으로 나눌 수 있다 (1). 겸상 적혈구 빈혈(sickle-cell anemia)이나 연골무형성증(achondroplasia)과 같은 멘델질환은 하나의 유전자가 결손되어 발생하고 멘델의 법칙에 따라 병이 전달된다. 반면, 당뇨, 암, 자폐증(autism) 및 류마티스관절염(rheumatoid arthritis; 이하 RA)과 같은 복합질환은 멘델의 유전 법칙을 따르지 않으며, 다양한 유전적 이상과 환경 요인의 복잡한

상호작용의 결과로 발생한다. 복합질환은 정해진 기준에 따라 하나의 이름 아래 분류되지만, 본질적으로 유사하거나 공통된 표현형을 나타내는 이질적인 환자군을 포함한다고 할 수 있다. 이러한 이질성(heterogeneity)은 크게 두 가지 요인으로 설명할 수 있다 (2). 하나는 조직내 세포 성분의 이질성이고, 또 하나는 환자에 따른 유전적 이질성이다. 세포 성분의 이질성은 질환의 속성에 따르지만, 유전적 이질성은 각 환자마다 같은 유전자에 이상이 발생하지 않을 수 있음을 의미한다.

<Received : December 6, 2013, Revised : December 10, 2013, Accepted : December 11, 2013>

Corresponding to : Wan-Uk Kim, Division of Rheumatology, Department of Internal Medicine, St. Vincent's Hospital, The Catholic University of Korea, 93, Jungbu-daero, Paldal-gu, Suwon 442-723, Korea. E-mail: wan725@catholic.ac.kr

pISSN: 2093-940X, eISSN: 2233-4718

Copyright © 2013 by The Korean College of Rheumatology

This is a Free Access article, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

RA는 나이와 성별에 따라 편향적으로 발생하면서 입체적인 상호작용을 보이는 여러가지 유전적 환경적 요인들의 변화를 포함하는 만성 진행성 자가면역 질환이다 (3). RA의 특징인 활막의 만성 염증은 관절에 손상을 주어 결국에는 변형과 기능장애를 유발한다. RA의 발병 기전에 대한 이해가 깊어지고 생명공학이 비약적으로 발전하면서 보다 효과적이고 안전한 새로운 약물들이 개발되어 임상에 도입되었으며, 지금도 많은 약물들이 개발 중이거나 임상시험 중이다. 특히, 종양괴사인자(tumor necrosis factor; TNF) 억제제는 RA 치료의 패러다임을 바꾸었는데, 임상 증상을 호전시켰을 뿐만 아니라 관절의 구조적 기능적 변형도 막을 수 있었다 (4,5). 하지만, 아쉽게도 항TNF 제제가 모든 환자에게 효과적인 것은 아니다. 항TNF 제제를 투여받은 환자의 약 30% 정도는 일정 수준의 임상적 호전을 보이지 못했고 (6-8), 양호한 반응을 보였던 일부 환자는 약물 치료 중 효과가 없어지거나 (9) 약물로 인한 부작용으로 약제를 중단하게 되었다. 뿐만 아니라 항TNF 제제 투여 후 관해에 도달했던 환자의 약 절반 정도는 1년 이상 그 상태를 유지하지 못했고 (10), 약물 중단 후에도 관해가 유지됐던 환자의 약 절반 정도는 재발하여 항류마티스약제(disease-modifying antirheumatic drugs; DMARDs)를 다시 복용하게 되었다 (11). 이런 실망스런 결과는 다른 생물학적 제제에서도 마찬가지였다 (12). 이는 생물학적 제제의 표적이 되는 TNF- α 나 interleukin-6 (IL-6)와 같은 사이토카인이 RA를 매개하는 핵심인자이지만 중간적 입장에 있을 뿐이며 이러한 사이토카인 한 두개를 조절하는 것만으로는 질병을 치료하는데 한계가 있음을 시사한다.

RA 환자들은 비슷한 증상과 소견을 가지고 있지만 서로 다른 임상 경과와 치료약물에 대한 차별적인 반응을 보인다. 그러므로 RA에서 병태생리학적으로 공통적인 변화를 일으키는 핵심적인 유전자 발굴함과 더불어 각 환자들의 유전적 특징에 따라 개별화된 치료 전략을 세울 수 있는 고유한 신호전달 경로 또는 표적 물질을 찾아내려는 노력이 필요한 상황이다.

질병에 대한 총체적 접근의 필요성

과거 10여 년간 유전 질환 연구는 대량 발굴(high-throughput) 글로벌 에세이(global assay) 기술과 대용량 데이터 분석법이 적용되면서 엄청난 변혁을 겪었다. 게놈 레벨 유전자 관련 분석(genome-wide association studies; GWAS)과 차세대 염기서열 시퀀싱(next generation sequencing; NGS), 질량분석기(mass spectrometry) 및 핵자기공명(nuclear magnetic resonance)과 같은 향상된 연구기법이 도입되면서 복합 질환의 특징을 결정하는 유전인자와 그 변이를 찾아 낼 수 있었다 (13,14). 이와 함께 다양한 질환에 대한 염기서열 변이와 각종 계층의 생체 분자 목록-예를 들어 전사체(mRNA), miRNA와 noncoding RNA, 단백질체(proteome), 대사체(metabolome) 등을 획득할 수 있었다(Figure 1). 이들

자료는 공용 데이터베이스나 출판물을 통해 공개되었고, 누구나 쉽게 접근하여 사용이 가능하다 (15).

전통적인 분자생물학적 연구법은 생체 분자의 신호 전달에 따른 특정 기능들, 예를 들면, 세포의 성장과 분화, 염증 반응과 같은 것들을 예측하여 작은 규모로 연구하는 것이다. 환원주의적 연구방법(reductionistic approach)이라고 하는 이 방법은 실제 특정 분자, 유전자, 또는 단백질을 분리하여 고유의 성장과 기능을 집중적으로 분석한다. 하지만, 살아있는 유기체 안에 있는 생체 분자는 개별적으로 작용하기 보다는, 다른 분자와 긴밀하게 협동하거나 복합체를 이루어 그 기능을 나타낸다. 뿐만 아니라, 대부분의 세포 내 구성성분은 다른 성분과 상호작용을 통해 기능을 수행하는데, 이러한 현상은 세포안, 또는 세포간, 심지어 기관 조직간에도 이루어진다 (16). 복합질환에서 분자간 세포간 비정상적 상호작용은 시간이 경과하고 해당 조직 또는 기관이 병적으로 변형되면서 그 정도와 범위가 변화한다. 따라서, 전통적인 생물학 연구법을 통해서서는 질환에서 확인된 단편적인 이상들을 조합하여 포괄적으로 이해하는데 한계가 있다.

이러한 한계를 극복하고 대량 발굴된 자료를 효율적으로 다루기 위해 시스템 수준의 접근방법이 필연적으로 대두하게 되었다. 지금도 그 정의, 개념 및 명명법에 다소 이견이 있지만 (17), 시스템 생물학(systems biology)은 기본적으로 대량 발굴 기술로 생산된 매우 복잡하고 이질적인 엄청난 양의 정보를 관리가능한 형태로 바꾸어 생물체의 현상과 기능을 이해하는 것을 목적으로 한다. 시스템 생물학을 이용하면 대규모 자료를 분석할 때 편향된 가설을 세우거나 편견을 가지지 않고 접근할 수 있으며, 응용통계학에 기반을 둔 과학적 자료분석기법을 바탕으로 데이터의 조직화된 모델을 구축하고 모델링을 통해 개별 자료를 객관적으로 바라볼 수 있다 (2). 미국 샌프란시스코 캘리포니아 대학교의 Ken Dill은 비행기 비유를 들어 시스템 생물학과 전통적인 생물학 연구방법의 차이점을 쉽게 설명하였다 (18). “비행기의 각 부분을 연구하여 그것들이 어떻게 작동

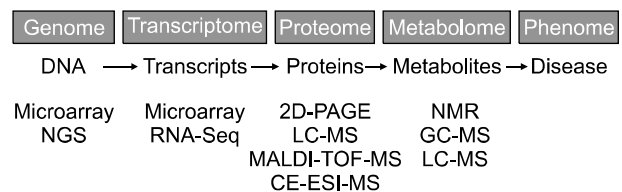


Figure 1. Systems genetics and techniques for analysis of complex disease. NGS, next generation sequencing; 2D-PAGE, two-dimensional polyacrylamide gel electrophoresis; MALDI-TOF-MS, matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry; CE-ESI-MS, capillary electrophoresis coupled with electrospray ionization mass spectrometry; NMR, nuclear magnetic resonance; GC-MS, gas chromatography-mass spectrometry; LC-MS, liquid chromatography-mass spectrometry.

하는지 기술할 수는 있지만, 그것들을 나열하는 것으로 비행기가 어떻게 하늘을 날 수 있는지 설명할 수 없다.” 생물학의 본질을 고려할 때, 생명 현상의 다면적 구조와 기능을 포괄하면서 많은 자료와 모델을 개념적으로 통합하기 위해서는 생물학, 컴퓨터 공학과 통계학의 공동작업을 통한 다학제간 기술 융합이 이루어져야 한다.

시스템 생물학을 기반으로 하는 연구법은 보통 5개 단계로 이루어진다 (19,20). 우선, 주어진 생물학적 질문에 대해 가장 적합한 시스템을 정한다. 시스템은 세포, 조직, 기관, 유기체뿐만 아니라 질병도 될 수 있다. 둘째, 연구하려고 하는 시스템의 구성 물질, 즉 mRNA 전사체, noncoding RNA, small interfering RNA, 단백질, 대사물질 및 기타 병태생리적 변수들 중 연구 목적에 부합하는 구성물질에 대해 글로벌 데이터를 생산한다. 셋째, 생산된 글로벌 데이터를 분석하여 차별적으로 발현된 유전자 혹은 단백질(differentially expressed genes or proteins)을 동정하고 연구 목적과 관련된 유전자군을 선별한다. 넷째, 시간 변화 또는 다양한 자극에 따라 동정된 구성 성분간의 상호작용이 어떻게 구성되는지 수리적으로 분석하여 네트워크 모델을 구축하고, 이를 생물학적으로 재해석함으로써 주어진 질문과 관련된 주요 네트워크 모듈을 예측한다. 마지막으로, 구성된 모듈이 대변하는 메커니즘을 적절한 조건에서 실험적으로 증명한다.

질병 측면에서 응용된 시스템 생물학을 시스템 또는 네트워크 의학(systems medicine or network medicine)이라고 하는데, 이는 질병에서 얻어진 대량 발굴 자료를 해석하기 위한 시스템 수준의 통합 접근법으로 최근 많은 주목을 받고 있다 (16,21). 시스템 의학은 질병의 증상, 예후, 치료에

대한 반응이 어떻게 조절되는지 파악할 수 있는 질환 관련 네트워크 모델을 제공할 뿐만 아니라 이를 통해 직접 질병의 분자적 복잡성을 이해할 수 있다 (16). 최근 들어 시스템 생물학이 더욱 발전하면서 대량 발굴된 생물학 정보를 편리하게 다룰 수 있는 컴퓨터 기술과 데이터 분석 기법이 개발되었으며, 전에 비해 보다 쉽게 이들 프로그램들을 사용할 수 있게 되었다 (2,17).

시스템 생물학의 연결체(interactome)와 네트워크

McKusick's Online Mendelian Inheritance in Man (OMIM[®]; <http://www.ncbi.nlm.nih.gov/omim>)은 지금까지 발표된 생물학의학관련 문헌들을 바탕으로 인간 유전자와 표현형을 집대성한 자료이다 (22). 현재까지 등록된 18,961개의 자료에서 약 10% 정도인 2,239개의 유전자가 질병을 유발하고, 3,770개의 질병에서 생체분자 이상이 있다고 정리하였다 (22). 이는 소수의 유전적 요인을 포함하면서 이들과 연결된 더 커다란 또는 복잡한 문제가 있음을 의미한다.

생물학적 활성이 나타나려면 단백질-단백질 사이, 단백질-DNA 사이, RNA 끼리 등 여러 계층의 분자간 상호작용이 적절하게 이루어져야 한다. 특히, 단백질-단백질간 상호작용(protein-protein interaction; PPI)은 시스템 수준에서 질병의 기전을 이해하는 기본적인 작용기작이다. 생체분자간 상호작용은 분자를 의미하는 마디(node 또는 vertex)와 마디 사이의 연결을 의미하는 선(link 또는 edge)으로 구성된 네트워크 도표로 표시된다. 상호작용의 형태에 따라 선은 방향성을 가질 수도 있다(Figure 2).

생물학적 상호작용을 표시하는 네트워크는 위상적 특징(topological property)이 있다 (23). 규칙을 표시에 반영하지

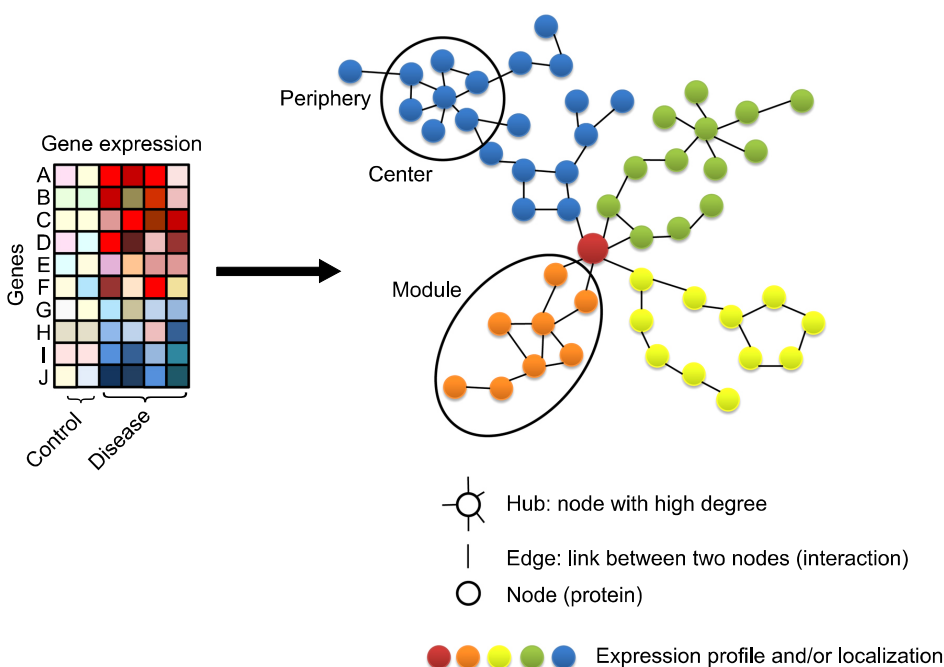


Figure 2. Transformation of data into network matrix and its components.

많은 생물학 네트워크 도표를 보면 연결된 선이 많은 마디와 그렇지 않은 마디로 구별된다. 연결된 선이 많은 마디를 중심(hub)이라고 하는데 (24), 생물학 기작에서 중요한 역할을 하고 물리적 기능적 연결을 표시한 네트워크에서 모듈(module) 구조를 이루는 근간이 된다(Figure 2) (25). 중심에 영향을 주는 유전자 변이(mutation)는 네트워크를 전체를 교란(perturbation)할 수 있지만, 주변부(periphery)에서 발생한 유전자 변이는 그 영향이 적다 (26). 네트워크의 연결성과 위상적 특징을 고려할 때, 질환과 관련된 유전자들을 연결된 집합 또는 모듈 형태로 다루는 것이 적절한 방법으로 간주되고 있다. 모듈은 형성된 네트워크 안에서 특정 세포 기능이 구현될 때 동원되는 유전자/단백질 또는 그 복합체의 집합을 의미한다 (27). 모듈은 기능 수행에 결정적인 마디 단백질을 중심으로 구성되고, 이는 관여된 구성성분이나 신호전달 경로, 시간에 따라 달라질 수 있다. 모듈을 구성하는 어느 하나 또는 이상의 분자에서 문제가 발생하여 핵심 모듈이 교란되면 질병이 발생하거나 병리적 특징이 결정될 수 있기 때문에 모듈의 개념은 매우 중요하다. 질병 모듈(disease module)의 개념은 “동일한 질환에서 관찰되는 유전자 또는 그 단백질의 이상은 서로 긴밀하게 상호작용할 가능성이 높다”라는 국소 가설(local hypothesis)을 근거로 한다 (16). 이러한 네트워크 기반 접근법과 모듈화는 실험실내 방법으로 구현하기 어려운 가상의 조건이나 교란 상태에서 생체 분자나 시스템의 행동을 예측할 수 있는 장점도 있다. 뿐만 아니라 전통적인 생물학 연구에서 해왔던 직관적 추론이나 문헌 고찰하는데 드는 수고를 덜고, 제시된 모듈내에서 예측가능 범위를 좁혀 예측함으로써 실험적 시행착오를 줄일 수도 있다.

다른 질환의 질병 모듈은 서로 겹칠 수 있고 한 질병 모듈에서 촉발된 교란은 다른 질병 모듈에 영향을 줄 수 있다. 따라서, 복합 질환들은 유사한 임상적 특징을 보이거나 동반질환(comorbidity)으로 나타날 수 있다 (28). 예를 들어, RA, 루푸스, 당뇨병은 전혀 다른 질병이지만 동맥경화증과 같은 심혈관 질환의 위험성이 유사한 정도로 높게 발생한다. 실제로, 질병간 네트워크와 질병-유전자 관련성이 확인된 유전자들을 시스템 접근법으로 분석한 연구에서 많은 질병들이 질병 네트워크 내에서 묶음 형태로 모여 있었는데, 이는 특정 이상 유전자들을 가진 질병들이 여러 가지 병적 기전을 공유하고 있다는 것을 시사한다 (29). 몇몇 약물들이 서로 다른 질환에서도 효과적인 것도 질병 모듈이 공유되고 연결성을 보이는 것으로 설명해 볼 수 있다 (30).

복잡하게 얽힌 네트워크에서 질병 관련 유전자 또는 단백질을 탐색하는 방법에는 크게 세가지가 있다 (16,31). 첫째, 연관법(linkage methods)이다. 질병 연관 간격(disease linkage interval)안에 존재하는 유전자에서 유래한 단백질은 기존에 알려진 질병 관련 단백질과 상호작용할 가능성이 높고, 그런 유전자가 잠재적인 질병 관련 유전자다 (32,33). 둘째, 질병 모듈법(disease module-based methods)이

다. 네트워크 모델에서 서로 연결되어 무리지어 있는 유전자들을 살펴보면 기능적으로 묶여있는 질병 모듈을 발견할 수 있고, 이 모듈을 구성하는 구성원이 잠재적인 질병 관련 유전자다 (32,34). 셋째, 확산법(diffusion-based methods)은 수리적으로 계산된 확률에 근거한 무작위 방문 알고리즘(random walker algorithm)을 따라 알려진 질병 관련 유전자에서 시작해서 다음 마디를 순차적으로 추적하는 것이다 (35,36). 이 알고리즘을 거치면 각 단백질마다 질병 연관 점수(disease-association score)가 매겨지는데, 이 점수로 그 유전자의 질병 연관성을 가늠할 수 있다.

류마티스관절염의 시스템적 접근

시스템 생물학 방법이 류마티스학 영역에도 도입되어 RA 이나 루푸스와 같은 질환에 대한 축적된 지식을 보다 잘 이해하고, 새로운 진단 표지자나 치료 표적을 발굴하는데 사용되기 시작하였다 (37,38). 하지만, 대부분 유전체 목록을 확보하여 비교하는 수준에 그치고 있고 진정한 의미의 시스템 접근법으로 통합 및 해석한 결과는 많지 않다.

RA 발생과 관련된 후보 유전자들과 우선 순위가 높은 차별적 발현 유전자들(differentially expressed genes, DEGs)이 대규모 GWAS와 전사체 연구를 통해 밝혀졌다 (39). RA 취약 유전자들(RA-susceptible genes)을 PPI 데이터베이스를 이용하여 분석하여 각 단백질 결과물간 물리적 상호작용을 보일 수 있는 연결고리를 밝히려는 시도가 있었는데, 이 네트워크는 RA 후보 유전자를 기반으로 작성되었고 중심이 되는 네트워크 마디와 부수적인 마디로 이루어져 있다. TRAF6, STAT1, IL2RB와 같은 이미 검증된 RA 취약 유전자들은 주요 마디에 위치하였다 (40). 흥미롭게도, 현재 효과적인 치료 표적인 CTLA4와 강력한 취약 유전자로 알려진 PTPN22는 기대했던 것만큼의 상호연결성을 보여주지 못했고, RA 취약성과 관련된 유전자 변이를 기능적 특징에 따라 재구성한 네트워크 분석에서는 오히려 주변부에 위치해 있었다. 이렇게 구성된 네트워크는 대량 발굴 자료와 PPI Human Protein Reference Database (HPRD)와 같은 데이터베이스를 이용하여 작성되었지만, 질병 관련 세포나 조직내에서 모듈의 역할과 모듈을 구성하는 단백질들 사이의 실제적인 상호작용들이 실험적으로 검증된 것은 아니었다.

온라인 데이터베이스를 탐색하여 GWAS 자료를 구해 다른 형태로 분석한 연구가 있다 (41). 이 연구에서는 RA 유전자 변이의 위험도와 예측도를 재평가하고, HPRD 데이터베이스를 이용해 PPI 네트워크를 구성하였다. PPI 네트워크 내에서 검증된 RA 취약 유전자들을 랜덤워크 리스타트(random walk with restart, RWR) 알고리즘을 이용해 근접도(proximity)를 고려하여 우선 순위를 정하였다. 이 알고리즘에서는 ZAP70이 가장 높은 순위였다. ZAP70 유전자 변이는 마우스에서 만성 자가면역 관절염을 일으킨다는 것이 확인된 바 있고 (42), 마우스 ZAP70 유전자 변이는

흉선 T세포 선택(thymic T-cell selection)에 영향을 주어 결국 RA 유사 관절염을 일으킨다고 설명하였다 (42). ZAP70은 HPRD 데이터베이스에 있는 RA 관련 유전자 단백질 중 PTPN22와 FCRL3와 직접 상호작용하는 것으로 나타났다 (41). 또, 이 연구자들은 EAGLE 알고리즘과 DAVID를 이용하여 RA 관련 네트워크에서 세가지 기능적 모듈을 찾아냈다. 백혈구 활성화와 분화(leukocyte activation and differentiation), 패턴인식 수용체 신호(pattern-recognition receptor signaling pathway), 케모카인과 그 수용체(chemokines and their receptors)가 여기에 해당한다. 비록 그 동안 발표된 자료들을 이용했지만, 이 결과는 시스템 생물학 방법이 질환과 관련된 다양한 유전적 변이들을 통합-분석함으로써 생물학적 신호경로를 이해하고 향후 관련된 연구의 방향을 잡는데 유용하다는 것을 보여주었다. 그러나 이 연구 역시 현재까지 잘 알려진 RA 분자적 발병기전의 일부만을 시스템적 접근법으로 재확인한 것에 그쳤다는 한계를 지닌다.

질병 모듈은 특정 질병 표현형과 연결되며, 기능 수행에 참여하는 여러 분자들이 교란된 상태-예를 들어 유전자 변이, 결손, 유전자 복제수 변이, 발현 차이 등등-로 모여있는 마디의 집합이다 (16). 질병 네트워크는 PPI내 특정 단백질의 역할을 교란 혹은 차단함에 의해 변화하고 재구성될 수 있다. 예를 들어 질병의 상태와 치료에 따라 질병 네트워크가 다르게 연결되거나 재구성될 수 있다. 본 연구팀은 활막조직, 말초혈액, 섬유모세포양 활막세포(fibroblast-like synovocyte; FLS)에서 얻어진 14개의 유전자 발현 자료를 선별하여 핵심적인 RA 관련 유전자를 찾아내고 RA에서 교란된 네트워크를 구성하였고 RA 진단과 치료에 유용한 잠재적 표적인자를 제시한 바 있다 (43). 우선, RA 교란 네트워크에서 RA FLS가 RA조직내 다양한 병리기작을 책임지는 핵심적인 세포임을 확인하였다. 둘째, 교란된 RA 네트워크가 항 TNF 치료에 의해 정상적인 수준으로 조정됨을 확인하였다. 흥미롭게도 B세포 활성화와 면역글로불린 모듈과 같은 B세포 관련 모듈은 항TNF 치료로 거의 변하지 않았는데, 이는 B세포 표적 치료가 항TNF 제제에 반응하지 않는 환자에서 효과적일 수 있음을 시사한다. 실제로, 항TNF 제제에 불응인 환자에서 항CD20 단일클론항체인 rituximab이 치료약으로 사용되고 있다 (44). 마지막으로 저자들은 RA 교란 네트워크로 균형이 깨진 유전자의 55%를 19개의 핵심 전사인자들이 조절한다는 것을 확인했다. 특히, RA 관련 세포 기작과 관련되어 있으나 진단 표지자나 치료 표적으로 한번도 보고된 적이 없는 후보 물질들을 발굴하였다. 이중, NFAT5는 고삼투압에서 활성화되는 삼투압보호 전사인자로 알려져 있었는데 RA 발병에 결정적인 역할을 한다는 것을 RA FLS와 NFAT5 반수결손 마우스에서 증명하여 발표한 바 있다 (45). 공개된 마이크로레이 자료를 이용해 RA 관련 기작간 상호작용 네트워크를 구성한 다른 연구결과에서도 NFAT5는 상호작용 네

트워크 사이의 모듈을 연결하는 중요한 전사인자로 밝혀졌다 (46). 그러므로 이와 같은 시스템 접근법은 RA 발병 기전에 대한 이해를 깊게 해 줄뿐 아니라 새로운 진단적 표지자와 치료 표적을 발굴하는데 유용한 역할을 할 수 있다는 것을 시사한다.

최근의 한 연구에서 TNF- α 가 조절하는 분자 네트워크를 확인하고 임상 지표와 관련된 치료반응의 다양성을 결정하는 분자 신호 특징들을 찾아내기 위해 RA 환자에서 항TNF 제제 투여 전후로 혈액을 채취한 후 유전자 발현 목록을 조사한 바 있다 (47). 그 결과, 22개의 전사체가 DAS28 점수 변동과 관련이 있는 것으로 확인되었고 이중 6개는 TNF 차단 치료의 대안이 될 수 있는 잠재적인 표적이었다. 특히, TNF 차단제 치료를 받은 RA 환자의 네트워크 분석에서 CD86이 압통 관절, 부종 관절 및 DAS28의 강력한 조절인자이며 TNF- α 와는 독립적인 조절기전으로 평가되었다. 이는 CTLA4-Ig인 abatacept를 이용해 CD86의 활성을 조절하는 것이 항TNF 제제의 대안치료가 될 수 있다는 임상결과와 일치하는 결과이다 (48).

시스템 연구의 한계와 대안

시스템 생물학은 대규모의 입체적 생체분자 자료를 해석하고 질병에서 유전형과 표현형 사이의 관련성을 총체적으로 이해하는 매우 유용한 연구방법이다. 하지만 몇 가지 한계점이 있다. 첫째, 세포 성분을 총괄적으로 측정하는 기술과 여기서 얻어진 데이터를 이용하여 세포의 기능적 네트워크 지도를 그릴 수 있는 데이터 분석 방법이 상당히 발전하였으나, 아직 일정 규모 이상의 비용이 들고 시간이 걸린다. 둘째, 사람 연결체(human interactome) 지도는 지금까지 발표된 생물학적 연구결과들을 기반으로 하고 있으나 이는 아직 완전한 형태가 아니다. 그리고 대량 발굴 자료의 분석을 통해 얻어진 모듈은 연결이 강할 수도 있지만 느슨할(sparse) 수도 있다. 이미 잘 알려진 질환 유전자들은 많이 연구되었기 때문에 더 많은 연결성을 갖게 되는 편향성이 있고, 현재의 PPI 자료에는 실제 기능적 연결고리가 누락된 것들이 많다. 셋째, 시스템 생물학 연구로 얻어진 네트워크 모델은 생화학이나 유전학의 전통적인 실험법을 대신할 수 없으며, 단지 자료를 통합하여 체계화한 원리/가설을 이끌어내는 역할을 할 뿐이다. 게다가, 연결성이 있다고 제시하나 실제 상호작용하는 성상에 대한 확실한 증거는 될 수 없다. 현재 사람 연결체에서 실험적으로 검증된 PPI는 130,000개로 평가되나, 상당수가 연결체 지도에서 아직 위치가 결정되지 못했다 (49). 실험적으로 검증되지 못했으나 사람 연결체에는 약 650,000개의 PPI가 존재할 것으로 추정되고 있다 (50). 따라서, 가상의 물리적 기능적 상호작용이 생물학적 의미를 획득하기 위해 실험적인 검증 과정을 거쳐야 한다. 넷째, 앞에서 언급한 마디와 연결은 조직(tissue)에 따라 구현되는 기능의 활성도가 다를 수 있으므로 조직 특이성 측면에서 다시 평가해야 한

다. 마지막으로, 지금까지 보고된 결과의 상당부분이 전사체 자료를 PPI 분석으로 해석한 것이다. 하지만, Altelaar 등에 따르면 mRNA 수준은 생체단백질 활성도 변화의 40% 정도만 설명해줄 수 있다 (51). 유전체는 약 25,000개의 유전자로 구성되지만 (52), 단백질은 백만개가 넘는 것으로 생각된다 (53). 전사(transcription) 또는 mRNA 수준에서의 변화는 유전자 대비 전사체 크기를 증가시키고, 수많은 번역후 수정(post-translational modification)은 유전체나 전사체 대비 단백질체의 복잡성을 기하급수적으로 증가시킨다 (51). 따라서, 질환에 따른 교란으로 결정되는 유전형과 표현형의 관련성을 정확하게 이해하기 위해서는 유전자부터 번역후 수정까지 전체 정보를 아우르는 시스템 분석이 필요하다.

마치는 말

RA 발병기전은 매우 복잡하고 이질적이어서 수 많은 세포 및 비세포 성분들이 상호 복잡하게 얽혀있고 영향을 주고 받는다 (54). 세포 성분은 염증 환경에서 유전적, 후생적(epigenetic) 변화를 겪고 비정상 세포로 전환되어 주변의 세포와 상호작용한다 (55). 특히, 유전형과 표현형은 유전체에 있는 정보만으로 알 수 없으며, 후생적 변화, 대체 스플라이싱(alternative splicing), 비번역 RNAs (non-coding RNAs), PPI 네트워크, 번역 후 수정을 포함하는 다양한 세포내 활동에 영향을 받는다 (56-59). 단백질들이 복합체를 형성하는 과정에서 혹은 시공간에 따라 변화하는 신호자극으로 인해 이러한 다양성은 더욱 복잡해진다. 따라서, 유전체 수준의 정보뿐만 아니라 전사체 및 단백질 수준의 자료를 모으고 이들을 임상조건에 맞는 조건과 치료에 따라 비교 분석하는 것이 필요하다. RA의 발병과정의 복잡성을 고려할 때, 관심을 갖는 특정 유전자 또는 단백질을 연구하는 전통적인 접근법으로 세포 안팎으로 얽혀있는 네트워크를 쉽게 이해할 수 없다. 또, 마우스와 같은 제한된 조건에서 실험적으로 검증된 내용은 실질적으로 많은 차이를 보이기 때문에 사람에게 적용하는데 한계가 있다 (60).

제한적이지만 현재까지 공개되어 있는 유전 질환과 연관된 약물 자료, 유전자 발현 정보 및 PPI를 이용하여 통합 분석했을 때, 놀랍게도 질환 유전자의 상당부분은 네트워크의 중심 위치에 있지 못하고 중심 단백질을 암호화하지 못하는 경향성을 보였으며, 그 발현 패턴은 네트워크의 주변부에 위치하고 있었다 (29). 이는 실제 질병을 유발하는 유전자와 교란된 모듈 또는 네트워크상에서 많은 연결성을 가져 다양한 표현형을 결정하는 핵심 유전자가 서로 다를 수 있음을 의미한다. 뿐만 아니라, 현재 사용되고 있는 많은 약물들은 질환 관련 단백질을 표적으로 하지 않고 네트워크의 이웃에 위치한 단백질을 표적으로 한다 (61). 즉, 상당수의 약물들이 질병을 완화시키는 수준이고, 질병의 근본적인 원인에 해당하는 교란된 단백질을 충분히 조절하지 못

하고 있는 것이다. 이는 앞서 언급한 RA를 시스템 수준으로 분석한 결과들에서도 제시되었듯이 (39-41,43,47), RA 치료를 위해 사용되고 있는 다양한 약제들이 왜 불충분한 효과를 보이고 있는지에 대해 설명하여 준다.

이러한 측면에서 질병과 의학을 시스템 생물학적 방법으로 재해석하고 풀어내는 것은 현재 우리가 질병 연구에서 직면하고 있는 문제들과 치료의 한계들을 총괄적으로 조망하고 해결의 방향성을 찾는 나침반과 같은 역할을 한다. 앞서 언급된 비행기 비유를 응용해보자. 비행기 엔진 본체에 이상이 생겨도 전원 버튼에 고장이 나도 비행기가 날지 못하는 결과는 동일하다. 하지만, 엔진 본체와 전원 버튼의 이상은 동일한 수준이 아니며 이를 수리할 때 접근하는 방식에 엄청난 차이가 있다. 즉, 관심을 갖는 유전자 또는 단백질 이상이 RA 발병 및 임상적 특징과 관련이 있는지, 있다면 이에 따른 세포의 기능적 변화가 어떤 것이고 그것의 파급 효과가 예측 가능한 범위인지, 예측한 범위 내에서 치료적 응용이 가능한 것인지를 확인해야 한다. 이를 위해서는 모듈 및 네트워크 수준에서 검토되어야 하며 그렇게 함으로 더욱 효과적이고 안전한 질병 치료 약물 개발이 가능할 것으로 생각된다.

현재 시스템 생물학의 기술적 발전 속도와 그에 따른 성과물들은 놀라운 결과들을 보여주었고 질병 연구의 새로운 기폭제가 되었다. 향후 전통적 환원주의적 생물학의 발달과 더불어 시스템 생물학의 보다 효율적이고 진보된 데이터 수집 및 분석 기술이 개발된다면 RA를 비롯한 다양한 류마티스 질환의 더 깊은 이해와 혁신적인 치료 기술 개발이 가능해질 것으로 기대한다.

References

1. Jin W, Qin P, Lou H, Jin L, Xu S. A systematic characterization of genes underlying both complex and Mendelian diseases. *Hum Mol Genet* 2012;21:1611-24.
2. Chuang HY, Hofree M, Ideker T. A decade of systems biology. *Annu Rev Cell Dev Biol* 2010;26:721-44.
3. Lee DM, Weinblatt ME. Rheumatoid arthritis. *Lancet* 2001;358:903-11.
4. Rubbert-Roth A, Finckh A. Treatment options in patients with rheumatoid arthritis failing initial TNF inhibitor therapy: a critical review. *Arthritis Res Ther* 2009;11 Suppl 1:S1.
5. Tanaka Y. Next stage of RA treatment: is TNF inhibitor-free remission a possible treatment goal? *Ann Rheum Dis* 2013;72 Suppl 2:ii124-7.
6. Lipsky PE, van der Heijde DM, St Clair EW, Furst DE, Breedveld FC, Kalden JR, et al; Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. Infliximab and methotrexate in the treatment of rheumatoid arthritis. *Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. N Engl J Med* 2000;343:1594-602.
7. Weinblatt ME, Keystone EC, Furst DE, Moreland LW, Weisman MH, Birbara CA, et al. Adalimumab, a fully

- human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum* 2003;48:35-45.
8. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;340:253-9.
 9. Finckh A, Simard JF, Gabay C, Guerne PA; SCQM physicians. Evidence for differential acquired drug resistance to anti-tumour necrosis factor agents in rheumatoid arthritis. *Ann Rheum Dis* 2006;65:746-52.
 10. Prince FH, Bykerk VP, Shadick NA, Lu B, Cui J, Frits M, et al. Sustained rheumatoid arthritis remission is uncommon in clinical practice. *Arthritis Res Ther* 2012;14:R68.
 11. Klarenbeek NB, van der Kooij SM, Güler-Yüksel M, van Groenendael JH, Han KH, Kerstens PJ, et al. Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. *Ann Rheum Dis* 2011;70:315-9.
 12. Aguilar-Lozano L, Castillo-Ortiz JD, Vargas-Serafin C, Morales-Torres J, Sanchez-Ortiz A, Sandoval-Castro C, et al. Sustained clinical remission and rate of relapse after tocilizumab withdrawal in patients with rheumatoid arthritis. *J Rheumatol* 2013;40:1069-73.
 13. Ku CS, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: where are we now? *J Hum Genet* 2010;55:195-206.
 14. Ormond KE, Wheeler MT, Hudgins L, Klein TE, Butte AJ, Altman RB, et al. Challenges in the clinical application of whole-genome sequencing. *Lancet* 2010;375:1749-51.
 15. Wang J, Zhang Y, Marian C, Ransom HW. Identification of aberrant pathways and network activities from high-throughput data. *Brief Bioinform* 2012;13:406-19.
 16. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;12:56-68.
 17. Likić VA, McConville MJ, Lithgow T, Bacic A. Systems biology: the next frontier for bioinformatics. *Adv Bioinformatics* 2010;268925.
 18. Vaughan C. The rise of systems biology. *UCSF Magazine* 2004;24.
 19. Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* 2001;2:343-72.
 20. MacLellan WR, Wang Y, Lusis AJ. Systems-based approaches to cardiovascular disease. *Nat Rev Cardiol* 2012;9:172-84.
 21. Furlong LI. Human diseases through the lens of network biology. *Trends Genet* 2013;29:150-9.
 22. Amberger J, Bocchini CA, Scott AF, Hamosh A. McKusick's Online Mendelian Inheritance in Man (OMIM). *Nucleic Acids Res* 2009;37:D793-6.
 23. Gursoy A, Keskin O, Nussinov R. Topological properties of protein interaction networks from a structural perspective. *Biochem Soc Trans* 2008;36:1398-403.
 24. Jeong H, Mason SP, Barabási AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature* 2001;411:41-2.
 25. Zotenko E, Mestre J, O'Leary DP, Przytycka TM. Why do hubs in the yeast protein interaction network tend to be essential: reexamining the connection between the network topology and essentiality. *PLoS Comput Biol* 2008;4:e1000140.
 26. Zhu X, Gerstein M, Snyder M. Getting connected: analysis and principles of biological networks. *Genes Dev* 2007;21:1010-24.
 27. Koutsogiannouli E, Papavassiliou AG, Papanikolaou NA. Complexity in cancer biology: is systems biology the answer? *Cancer Med* 2013;2:164-77.
 28. Bauer-Mehren A, Bundschuh M, Rautschka M, Mayer MA, Sanz F, Furlong LI. Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases. *PLoS One* 2011;6:e20284.
 29. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. *Proc Natl Acad Sci U S A* 2007;104:8685-90.
 30. Suthram S, Dudley JT, Chiang AP, Chen R, Hastie TJ, Butte AJ. Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. *PLoS Comput Biol* 2010;6:e1000662.
 31. Cho DY, Kim YA, Przytycka TM. Chapter 5: Network biology approach to complex diseases. *PLoS Comput Biol* 2012;8:e1002820.
 32. Navlakha S, Kingsford C. The power of protein interaction networks for associating genes with diseases. *Bioinformatics* 2010;26:1057-63.
 33. Oti M, Snel B, Huynen MA, Brunner HG. Predicting disease genes using protein-protein interactions. *J Med Genet* 2006;43:691-8.
 34. Wu X, Jiang R, Zhang MQ, Li S. Network-based global inference of human disease genes. *Mol Syst Biol* 2008;4:189.
 35. Köhler S, Bauer S, Horn D, Robinson PN. Walking the interactome for prioritization of candidate disease genes. *Am J Hum Genet* 2008;82:949-58.
 36. Vanunu O, Magger O, Ruppin E, Shlomi T, Sharan R. Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol* 2010;6:e1000641.
 37. Chiche L, Jourde-Chiche N, Pascual V, Chaussabel D. Current perspectives on systems immunology approaches to rheumatic diseases. *Arthritis Rheum* 2013;65:1407-17.
 38. Sirota M, Butte AJ. The role of bioinformatics in studying rheumatic and autoimmune disorders. *Nat Rev Rheumatol* 2011;7:489-94.
 39. Toonen EJ, Barrera P, Radstake TR, van Riel PL, Scheffer H, Franke B, et al. Gene expression profiling in rheumatoid arthritis: current concepts and future directions. *Ann Rheum Dis* 2008;67:1663-9.
 40. Viatte S, Plant D, Raychaudhuri S. Genetics and epi-

- genetics of rheumatoid arthritis. *Nat Rev Rheumatol* 2013;9:141-53.
41. Nakaoka H, Cui T, Tajima A, Oka A, Mitsunaga S, Kashiwase K, et al. A systems genetics approach provides a bridge from discovered genetic variants to biological pathways in rheumatoid arthritis. *PLoS One* 2011;6:e25389.
 42. Sakaguchi N, Takahashi T, Hata H, Nomura T, Tagami T, Yamazaki S, et al. Altered thymic T-cell selection due to a mutation of the ZAP-70 gene causes autoimmune arthritis in mice. *Nature* 2003;426:454-60.
 43. You S, Cho CS, Lee I, Hood L, Hwang D, Kim WU. A systems approach to rheumatoid arthritis. *PLoS One* 2012;7:e51508.
 44. Keystone E, Burmester GR, Furie R, Loveless JE, Emery P, Kremer J, et al. Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. *Arthritis Rheum* 2008;59:785-93.
 45. Yoon HJ, You S, Yoo SA, Kim NH, Kwon HM, Yoon CH, et al. NF-AT5 is a critical regulator of inflammatory arthritis. *Arthritis Rheum* 2011;63:1843-52.
 46. Wu G, Zhu L, Dent JE, Nardini C. A comprehensive molecular interaction map for rheumatoid arthritis. *PLoS One* 2010;5:e10137.
 47. Xing H, McDonagh PD, Bienkowska J, Cashorali T, Runge K, Miller RE, et al. Causal modeling using network ensemble simulations of genetic and gene expression data predicts genes involved in rheumatoid arthritis. *PLoS Comput Biol* 2011;7:e1001105.
 48. Schiff M. Abatacept treatment for rheumatoid arthritis. *Rheumatology (Oxford)* 2011;50:437-49.
 49. Venkatesan K, Rual JF, Vazquez A, Stelzl U, Lemmens I, Hirozane-Kishikawa T, et al. An empirical framework for binary interactome mapping. *Nat Methods* 2009;6:83-90.
 50. Stumpf MP, Thorne T, de Silva E, Stewart R, An HJ, Lappe M, et al. Estimating the size of the human interactome. *Proc Natl Acad Sci U S A* 2008;105:6959-64.
 51. Altelaar AF, Munoz J, Heck AJ. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat Rev Genet* 2013;14:35-48.
 52. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004;431:931-45.
 53. Jensen ON. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* 2004;8:33-41.
 54. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *N Engl J Med* 2011;365:2205-19.
 55. Bottini N, Firestein GS. Duality of fibroblast-like synoviocytes in RA: passive responders and imprinted aggressors. *Nat Rev Rheumatol* 2013;9:24-33.
 56. Frenkel-Morgenstern M, Lacroix V, Ezkurdia I, Levin Y, Gabashvili A, Prilusky J, et al. Chimeras taking shape: potential functions of proteins encoded by chimeric RNA transcripts. *Genome Res* 2012;22:1231-42.
 57. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 2011;146:353-8.
 58. Baker M. Proteomics: The interaction map. *Nature* 2012;484:271-5.
 59. Bensimon A, Heck AJ, Aebersold R. Mass spectrometry-based proteomics and network biology. *Annu Rev Biochem* 2012;81:379-405.
 60. Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, et al; Inflammation and Host Response to Injury, Large Scale Collaborative Research Program. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci U S A* 2013;110:3507-12.
 61. Yildirim MA, Goh KI, Cusick ME, Barabási AL, Vidal M. Drug-target network. *Nat Biotechnol* 2007;25:1119-26.