

---

# 고등학교 과학 및 생물교과서 과학용어 네트워크 분석

## Analysis of Scientific Item Networks from Science and Biology Textbooks

---

박별나, 이윤경, 구자을, 홍영수, 김학용  
충북대학교 자연과학대학 생화학과

Byeol Na Park(byeolpark@cbnu.ac.kr), Yoon Kyeong Lee(biokorea01@cbnu.ac.kr),  
Ja Eul Ku(chaeul@cbnu.ac.kr), Young Soo Hong(younghong@cbnu.ac.kr),  
Hak Yong Kim(hykim@cbnu.ac.kr)

---

### 요약

교과서에 쓰인 과학 용어 네트워크를 구축하여, 네트워크의 구조, 관련 정보 및 연관 관계를 분석하기 위하여 핵심용어를 도출하였다. 본 연구에서는 과학, 생물1 및 생물2 교과서 각 과목별로 출판사 세 곳을 선정하고 각각의 교과서에서 추출한 용어들을 노드로, 한 문장 안에 있는 과학 용어를 링크로 연결하여 네트워크를 구축하였다. 모든 교과서의 과학 용어 네트워크는 척도 없는(scale-free) 네트워크의 특성을 보여 주었다. 복잡한 (complex) 네트워크에서 가중치가 낮은 것부터 제거하는 방법인 k-core 알고리즘을 적용하여 핵심 (core) 네트워크를 구축하였는데, 몇 개의 모듈이 연결되는 형태를 보여주었다. 과학교과서의 경우에는 물리, 화학, 생물, 지구과학 분야별로 크게 네 개의 모듈을 형성하였고, 생물1과 생물2 교과서는 각 단원별로 용어들이 모여 있는 특성을 지닌 네트워크를 나타냈다. 본 연구에서 복잡한 네트워크에서 핵심 네트워크를 구축하여 유용한 정보를 도출할 수 있는 가능성을 제시하였다.

■ 중심어 : | 생물 교과서 네트워크 | 과학 교과서 네트워크 | k-core 알고리즘 |

### Abstract

We extracted core terms by constructing scientific item networks from textbooks, analyzing their structures, and investigating the connected information and their relationships. For this research, we chose three high-school textbooks from different publishers for each three subjects, i.e., Science, Biology I and Biology II, to construct networks by linking scientific items in each sentence, where used items were regarded as nodes. Scientific item networks from all textbooks showed scale-free character. When core networks were established by applying k-core algorithm which is one of generally used methods for removing lesser weighted nodes and links from complex network, they showed the modular structure. Science textbooks formed four main modules of physics, chemistry, biology and earth science, while Biology I and Biology II textbooks revealed core networks composed of more detailed specific items in each field. These findings demonstrate the structural characteristics of networks in textbooks, and suggest core scientific items helpful for students' understanding of concept in Science and Biology.

■ keyword : | Biology Textbook Network | Science Textbook Network | K-core Algorithm |

---

\* 본 연구는 2008학년도 충북대학교 학술연구지원사업의 연구비 지원에 의하여 수행되었다.

접수번호 : #100217-006

심사완료일 : 2010년 03월 17일

접수일자 : 2010년 02월 17일

교신저자 : 김학용, e-mail : hykim@cbnu.ac.kr

## I. 서론

현대사회로 나아가면서 과학의 중요성이 점차 높아지고 있다. 따라서 학생들이 과학에 대해서 올바른 개념을 정립시키고, 흥미를 유발시키고, 나아가서 폭넓게 탐구할 수 있도록 하는 일은 매우 중요하다. 현재 학생들이 공부하는데 가장 쉽게 접하는 자료는 학교에서 사용하는 교과서이다. 교과서들을 출판사별로 또는 다른 나라의 교과서나 교육과정의 각 차수에 따라서 내용을 비교, 분석을 하는데 각 단원이 교과서 내에서 차지하고 있는 분량을 페이지수를 사용하여 비교하는 방법이나 단원들이 추가되거나 삭제되는 것을 비교하는 방법을 이용하여 좀 더 나은 교과서를 만들고자 하는 노력을 하고 있다[1].

네트워크는 노드라고 부르는 점과 그들 사이를 연결하는 링크(연결선)로 이루어진 하나의 기하학적 구조이다. 자연현상이나 사회현상에서도 이들 현상들을 구성하고 있는 요소들과 관계를 보여주는 네트워크에 관한 연구가 최근에 많이 이루어지고 있다. 사회 네트워크[2-4], 생명체 및 인간 질병 네트워크[5-8], 인터넷 네트워크[9], 소셜 속 등장 인물네트워크[10-12], 및 물리 지식 네트워크[13] 등 사회 및 자연 전 분야에서 보여주는 네트워크의 구조적 성질을 분석하여 의미를 도출하는 연구가 확산되고 있다.

따라서 네트워크에 대한 개념을 교과서에 적용하여, 단순히 단원의 유무나 분량으로 비교하는 것이 아니고 교과서를 서술하는데 사용된 용어들 간의 상관관계를 분석해봄으로서 학생들이 공부하는데 도움이 되는 용어들을 선별할 수 있다. 나아가서 교과서가 어떤 방향으로 서술되어야 효과적으로 개념을 습득하는데 도움을 주는지 까지를 제시해 줄 수 있을 것으로 기대된다.

본 논문에서는 과학, 생물1, 생물2 교과서에 쓰인 과학 용어들을 추출하여, 이 용어를 노드로 하고 한 문장 내에 함께 사용된 과학 용어를 연결하는 링크로 하여 과학용어 네트워크를 구축하였다. 구축한 네트워크를 통해서 구조를 분석하여 과학교과서의 서술방법을 알아 볼 수 있으며, 전체적으로 다양하게 과학을 접할 수 있도록 개념을 형성해 주는데 쓰이는 용어들의 적절성

도 알아보았다. 또한 중요하다고 생각되는 핵심(core) 과학 용어를 도출하였는데, 이 용어들은 앞으로 교과서를 서술하는데 있어서 참고할 수 있는 중요한 정보로 활용할 수 있을 것이다.

## II. 연구자료 및 연구 방법

### 1. 과학용어 데이터베이스 구축

본 연구에서는 분석에 쓰일 용어를 추출하는데 고등학교 과정에서 배우는 과학, 생물1, 생물2 교과서를 대상으로 하였다. 각 과목별로 7차 교육과정에 의거해 출판된 교과서 8종 중 임의로 3종을 정하여 사용하였다. 과학은 금성, 디딤돌, 중앙 출판사 교과서를, 생물1 및 2는 금성, 대한, 중앙 출판사 교과서를 각각 사용하였다.

과학, 생물1, 생물2 교과서에서 용어를 선택한 기준은 다음과 같이 정하였다. ① 교과서의 본문 중에 서술된 명사만 선택한다. ② 띄어쓰기의 경우 나열된 2개 이상의 명사가 합쳐져서 하나의 특정한 의미를 나타내면 띄어서 쓰지 않고 붙여서 하나의 용어로 만든다. ③ 한 문장 내에 같은 용어가 반복되면 한 번만 입력한다. ④ 반복되는 용어를 생략하여 표기한 경우 생략하기 이전의 용어로 간주한다. ⑤ 괄호 안에 표시된 용어도 하나의 독립 용어로 추출한다. ⑥ 독립된 영문 단어 및 영문 단어가 포함된 용어도 하나의 용어로 관주한다. ⑦ 괄호 안에 쓰인 바로 앞의 용어를 의미하는 영어약자는 제외한다. ⑧ 아라비아 숫자는 한글로 전환하여 사용한다. ⑨ 한 문장에서 추출된 용어가 하나일지라도 버리지 않고 사용하여 다른 문장에서 나온 용어가 있다면 연결될 수 있도록 한다.

일차적으로 위에서 제시한 기준으로 용어를 추출한 다음, 2005년에 한국과학기술단체총연합회에서 발간한 과학기술대사전에 등재된 용어만을 다시 선별하였다.

### 2. 네트워크 구축

네트워크의 구축은 복잡계 네트워크 분석 프로그램 파이엑(Pajek)(<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>)을 사용하였다. 네트워크로 만들기 위해서 한 문

장 내에서 함께 나온 과학 용어는 서로 연관이 있다고 간주하고 각각의 용어를 서로 연결 (상호작용)해 주었다. 네트워크에서 하나의 용어가 다른 하나의 용어와 상호작용을 하게 되면 그 링크가 하나의 선을 이루도록 하였고, 하나의 용어가 어떤 다른 용어와도 상호작용하지 않으면 이 링크에 대해서는 연결선이 그려지지 않게 하는 방법으로 네트워크를 구축하였다.

### 3. 복잡계 네트워크 분석

그래프 이론에서 복잡한 그래프를 단순화 시키는 방법인 k-core 알고리즘[14]을 바탕으로 만든 파이엑(Pajek)프로그램을 사용하여 복잡한 네트워크를 분석하였다. K-core 알고리즘은 네트워크의 노드들이 k 이상의 링크수를 가지는 노드들로 구성된 네트워크로, 이 네트워크가 형성될 때까지 링크수가 k 이하인 노드를 제거하는 과정을 반복하게 된다. 그러므로 전체 네트워크에서 k 값이 작은 노드일수록 가중치가 낮아 먼저 제거된다. 따라서 과학, 생물1, 생물2 교과서 모두 파이엑 프로그램을 사용하여 k-core 값을 증가시켜가면서 네트워크의 크기를 줄여나갔다. 가중치가 낮은 용어가 제거되고 남은 가중치가 높은 핵심용어 분포를 구분하기 위해 싸이토스케이프(Cytoscape) 프로그램 (<http://www.cytoscape.org/>)을 사용하여 시각화 하였다.

### III. 연구결과 및 분석

과목별, 출판사별로 추출해 낸 용어들 간의 네트워크는 대부분의 용어들이 서로 연결되어있는 큰 네트워크와 몇 개의 소수 용어들이 떨어져있는 아주 작은 네트워크로 되어있다. 따라서 전체 네트워크에서 상호작용하는 모든 용어들 중 가장 큰 네트워크를 형성하지 않고 부분적으로 몇 개 연결되어 있는 일부 용어들은 제거하였다. 이들을 제거하고 만든 네트워크를 largest 네트워크라고 명명하였고, 본 논문에서는 largest 네트워크에서 사용된 용어들을 분석을 하였다[표 1].

표 1. 과학, 생물1, 생물2 과학용어 네트워크 특성 분석.

과목	출판사	Largest Network			
		N	L	<K>	<C>
과학	금성	780	3,401	4.36	0.73
	디딤돌	840	3,925	4.67	0.71
	중앙	811	3,368	4.15	0.75
생물1	금성	1,111	6,630	5.97	0.73
	대한	1,229	7,861	6.40	0.76
	중앙	1,326	7,397	5.58	0.73
생물2	금성	1,912	11,660	6.10	0.72
	대한	1,418	8,627	6.08	0.75
	중앙	1,519	10,174	6.70	0.72

N: 용어수, L: 연결수, <K>: 평균 연결계수, <C>: 평균 뭉침계수.

먼저 세 과목 교과서를 구성하고 있는 네트워크의 노드(과학용어) 수와 평균 연결계수는 과학에서 생물1, 생물2로 갈수록 증가하는 것을 보여주고 있다. 이는 과학은 초보적 과학지식을 다루는 것에 반해, 생물1은 과학에 비해 전문적인 생물 용어를 다루고 생물2에서는 생물1에 비해 더 전문적인 용어를 다루고 있기 때문에 용어의 수가 증가하고 이들 용어 간의 상관관계도 더 복잡하기 때문인 것으로 판단된다. 이에 비해 뭉침계수는 비슷하게 나타났는데, 각 과목의 용어 사이의 뭉침 정도는 과목과 상관없이 비슷한 것을 반영하고 있다.

앞에서 살펴본 바와 같이 각 과목 3종의 교과서의 네트워크 특성이 비교적 비슷하고 앞으로 진행될 연구 결과도 비슷하기 때문에, 우리는 과학은 디딤돌 교과서, 생물1은 중앙교과서, 생물2는 대한교과서를 중심으로 알아보려고 한다. 특별한 언급이 없다면 다른 모든 교과서도 같은 특성을 보여준다는 것을 의미하며 연구 진행 중 상이한 부분이 있으면 그 시점에 언급할 것이다. [그림 1]에서 840개의 과학용어와 3,925개의 링크로 구성되어 있는 과학 디딤돌 교과서 네트워크를 보여주고 있다. 다른 모든 교과서 네트워크도 유사한 복잡계 네트워크를 형성하고 있음을 확인하였다.

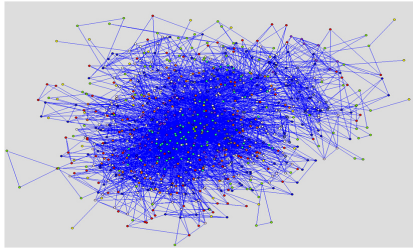


그림 1. 과학 디딤돌 교과서 largest 네트워크.

교과서 네트워크도 사회네트워크[15]나 소설에 등장하는 인물 네트워크[12]와 마찬가지로 척도 없는 네트워크의 특성을 보여주는지를 확인하기 위하여 연결수 누적 분포 함수를 측정하였다[그림 2]. 그림에서 x축은 연결수를 보여주며 y축은 인물 누적 분포수를 의미한다. 비록 [그림 2]에서 연결 수가 1과 2인 곳에서의 P(K)값이 다소 낮은 경향은 있으나 선형회귀분석(linear regression)에서 R2값(0.76 ~ 0.81)이 유의하기 때문에 멱함수 분포를 따르는 것으로 볼 수 있으며, 따라서 척도 없는 네트워크의 특성을 보여주고 있다고 할 수 있다[그림 2].

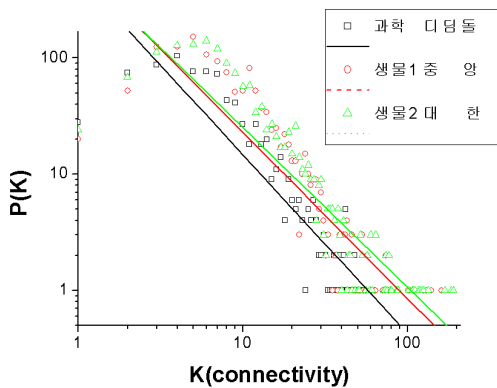


그림 2. 과학, 생물, 및 생물2 교과서 네트워크의 멱함수 분포.

과학에서 생물2로 갈수록 connectivity 값(x 절편값)이 증가하는 경향을 볼 수 있는데, 과학보다는 생물에 관한 기술이 좀 더 심층적이기 때문에 복잡한 연결 관계를 맺고 있음을 보여준다. 이는 앞에서 언급한 용어의 수가 증가하고 상관관계가 더 복잡해지는 것과 일치

하고 있다[표 1].

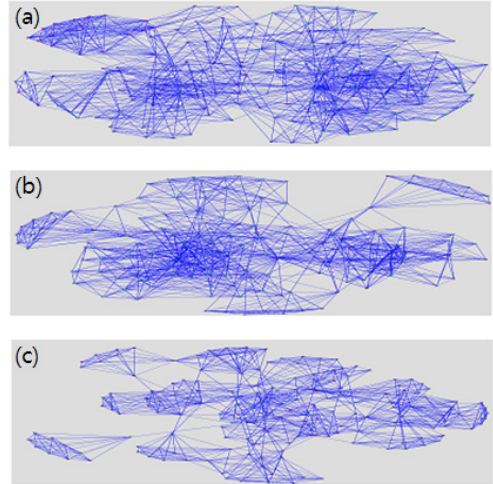


그림 3. 과학 교과서 core 네트워크. (a)금성 k-core =7, (b)디딤돌 k-core =8, (c)중앙 k-core =7.

복잡계 네트워크의 구조적 특성에 대한 연구는 지난 10여 년 동안 다양한 분야에서 급속도로 증가하는 양상을 보여 왔는데, 최근 들어 이 분야에 대한 연구가 완만하게 감소하는 경향이 있다. 이는 복잡계 네트워크의 구조적 분석을 통해서 유용한 정보를 얻었지만 구조적 분석에서 얻을 수 있는 정보의 한계가 있음을 반영한 결과이다. 따라서 복잡계 네트워크로부터 유용한 정보를 얻기 위한 시도와 접근이 요구되는 시점이다.

본 연구에서는 복잡한 네트워크 중에서 비교적 가치가 낮은 연결고리를 단계별로 제거시킴으로서 단순하지만 중요한 핵심(core) 네트워크를 구축하고자 하였다. 이를 위해서 Ignacio[14]등이 개발한 k-core 알고리즘을 바탕으로 구축한 파이엑(Pajek) 프로그램을 이용하여 k값이 1부터 차례로 적용하여 각 교과서의 핵심 네트워크를 구축하였다[그림 3-5]. 과학, 생물1, 및 생물2 교과서 모두 k-core값이 증가함에 따라 과학용어들이 서로 더욱 뭉쳐서 모듈(module)을 형성하는 경향을 보여주었다[그림 3-5]. 이 k값이 특정 값 이상이 되면 모듈 구조가 파괴되어 단순한 덩어리 네트워크로 남기 때문에, 그 이전 k값을 최적으로 삼아 k-core 값으로 명명

하였다. 과학 교과서의 경우, 금성은 k값이 7에서, 디딤돌은 k값이 8에서, 중앙은 k값이 7에서 비교적 명확한 모듈 구조를 보여주었다[그림 3].

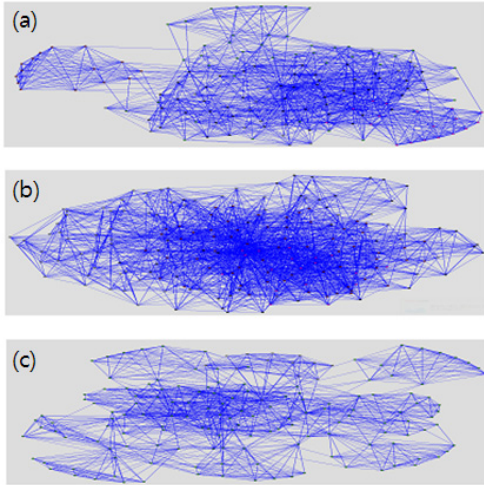


그림 4. 생물1 교과서 core 네트워크. (a)금성 k-core =10.(b)대한 k-core =11, (c)중앙 k-core =10.

생물1 교과서의 경우, 금성은 k값이 10에서, 대한은 k값이 11에서, 중앙은 k값이 10에서 비교적 명확한 모듈 구조를 보여주었다[그림 4]. 그러나 금성이나 중앙에 비해 대한 교과서는 모듈 구조가 명확하지 않았다. 대한 교과서의 경우 모듈 구조를 찾기 위해 k 값을 지속적으로 증가시켰으나 모듈 구조를 나타내지 않고 네트워크 덩어리만이 축소되는 경향을 보여주었다. 이는 교과서 집필 당시, 생물 용어 사이의 구분을 명확하고 체계적으로 분류하기 보다는 한 소단위 분야의 용어와 다른 소단위 분야의 용어를 상호 혼합하여 설명하였기 때문인 것으로 풀이된다.

생물2 교과서의 경우, 금성은 k값이 12에서, 대한은 k값이 12에서, 중앙은 k값이 13에서 비교적 명확한 모듈 구조를 보여주었다[그림 5]. 생물2의 경우는 오히려 금성이나 중앙에 비해 대한 교과서가 더 명확한 모듈 구조를 보여주고 있다. 이것 역시 앞에서 언급한 집필과정에서의 성향이 반영된 것으로 볼 수 있다.

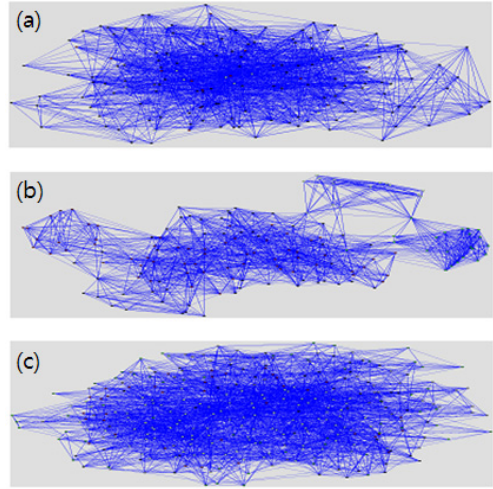


그림 5. 생물2 교과서 core 네트워크. (a)금성 k-core =12, (b)대한 k-core =12, (c)중앙 k-core =13.

K-core 알고리즘은 가중치가 낮은 링크를 제거하기 때문에 남아있는 네트워크의 핵심용어는 가중치가 매우 높은 중심 과학용어로 볼 수 있다. 따라서 교과서 핵심 네트워크에서 핵심 과학용어 네트워크의 노드에 실제 과학용어를 부여하여 교과서 핵심 과학용어 네트워크로 전환한 후 핵심 과학용어를 추출하였다. 과학 금성교과서 핵심용어 178개, 과학 디딤돌교과서 핵심용어 147개, 과학 중앙 교과서 핵심용어 180개를 추출하였으며, 생물1 금성교과서 핵심용어 142개, 생물1 대한 교과서 핵심용어 180개, 생물1 중앙교과서 핵심용어 157개를, 생물2 금성교과서 핵심용어 156개, 생물2 대한교과서 핵심용어 151개, 생물2 중앙교과서 핵심용어 146개를 각각 추출하였다[부록]. 이들은 과학과 생물을 이해하는데 핵심 용어로서 중요한 자료로 활용할 수 있을 것이다. 앞에서 언급한 바와 같이 이들 중 과목별 각 교과서에서 비교적 모듈구조를 잘 보여 주고 있는 과학 디딤돌 교과서[그림 6], 생물1 중앙 교과서[그림 7]와 생물2 대한 교과서[그림 8]의 핵심 과학용어 네트워크로 전환하였다.

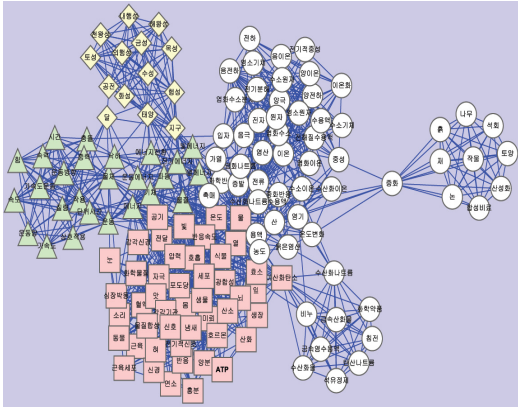


그림 6. 과학 디딤돌 교과서 핵심 과학용어 네트워크. (k-core =8)

과학 디딤돌 교과서의 핵심 과학 용어는 물리, 화학, 생물, 지구과학으로 정확하게 분리되는 모듈 구조를 보여 주고 있다[그림 6].

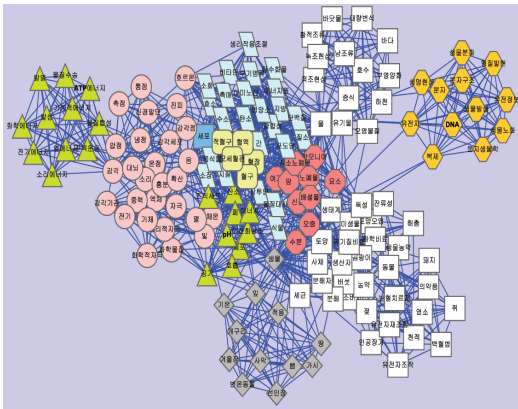


그림 7. 생물1 중앙 교과서 핵심 과학용어 네트워크. (k-core =10)

생물1과 생물2에서도 비슷한 양상을 보여주고 있는데, 이 네트워크에서는 교과서 각 단위별로 모듈을 형성하는 경향성을 보여주고 있다. 그러나 과학교과서와는 달리 일부 용어들이 모듈 사이를 연결하는 특성을 보여주고 있는데, 이들 연결고리 역할을 하는 용어들은 각 단위의 과학적 설명을 이어주는 용어로 판단된다. 예를 들어, [그림 8]에서 ‘돌연변이’나 ‘체세포’는 진화 단위와 분자 생물학적 기작 및 형질전환과 같은 생명공

학 단위에서도 널리 사용되기 때문에 모듈을 이어주는 연결고리로서 역할을 하고 있음을 보여주고 있다.

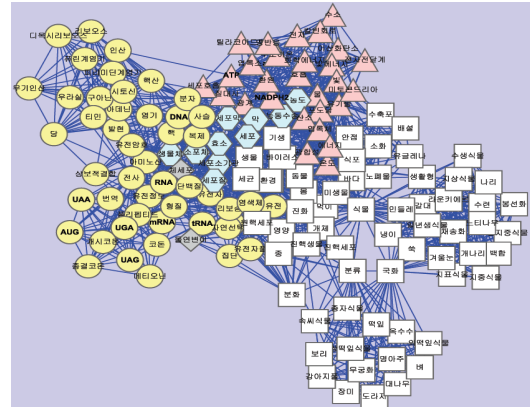


그림 8. 생물2 대한 교과서 핵심 과학용어 네트워크. (k-core =12)

본 연구에서 조사한 9권의 교과서로부터 얻은 핵심 과학용어 중에서 링크수가 가장 많은 상위 10개 핵심 과학 용어 목록을 [표 2]에 제시하였다. 과학 교과서에서는 물 또는 몸이 가장 링크수가 많은 것으로 나타났는데, 물과 몸은 생물뿐만 아니라 화학에서도 자주 사용되는 용어이기 때문인 것으로 풀이된다. 금성이나 디딤돌에서는 물리, 화학, 생물, 지구과학 용어가 고르게 분포되어 있으나 중앙에서는 생물 용어가 대부분이며 화학 일부를 차지하고 있다. 과학교과서를 분석해보면 중앙 교과서가 생명을 다룬 부분은 다른 교과서와 유사하나 유독 생물과 화학 용어를 많이 사용하는 환경 분야에 대해서는 디딤돌 교과서에 비해 1.5배 이상을 다루고 있다. 비록 금성 교과서도 환경 분야를 비중 있게 다루었지만 주로 대기와 물의 오염에 관한 용어인 물, 대기, 지구 등이 많았다. 따라서 중앙 교과서의 상위 10개 핵심 용어가 생물과 화학 용어가 많이 나타나는 원인이다.

일반적으로 생물1은 생태계, 몸과 기관 수준에 대해 많이 다루고 있는데 반해, 생물2는 세포 수준이나 분자 수준인 DNA나 단백질, 진화와 광합성, 생명공학과 기술 분야를 다루고 있다. 생물교과서의 핵심 과학 용어에서도 이러한 사실이 그대로 반영되고 있다. 생물1 교과서 3종 모두는 핵심 과학 용어들이 몸, 혈관과 같은

표 2. 과학, 생물1, 생물2 교과서 상위 10개 핵심용어 목록.

과학 교과서			
	금성	디딤돌	중앙
1	물	물	몸
2	빛	에너지	물
3	물질	빛	전달
4	식물	온도	빛
5	지구	물질	산
6	공기	염산	생물
7	에너지	몸	빛물
8	전해질	뇌	자극
9	대기	물체	피부
10	전류	전류	물질
생물1 교과서			
	금성	대한	중앙
1	몸	몸	물
2	혈액	물질	몸
3	모세혈관	혈액	생물
4	물	세포	산소
5	영양소	산소	단백질
6	포도당	조직세포	무기염류
7	세포	영양소	이산화탄소
8	이산화탄소	호르몬	영양소
9	무기염류	이산화탄소	혈액
10	노폐물	생물	감각
생물2 교과서			
	금성	대한	중앙
1	광합성	식물	생물
2	물	생물	동물
3	식물	단백질	식물
4	생물	세포	물
5	세포	분자	세포
6	세균	DNA	광합성
7	물질	광합성	몸
8	분류	효소	단백질
9	산소	물	유기물
10	에너지	에너지	에너지

핵심 과학 용어 네트워크에서 각 과목별 교과서에서 링크수가 가장 많은 상위 10개 용어의 목록.

기관을 나타내는 용어 및 생물에서 나타나는 일반적인 용어들이 많이 나타난다[표 2]. 생물2는 분자나 세포 수준을 보여주는 핵심용어들이 나타나는데, DNA, 유전자, 단백질, 효소 등이며, 이외에도 세포, 광합성, 광합

성의 핵심어인 에너지가 상위 10위 핵심 과학 용어에 포함되어 있다. 핵심 용어 이외에도 앞에서 언급한 과학의 160여 핵심 용어, 생물1의 160여 핵심용어, 생물2의 150여 핵심 용어는 과학과 생물을 이해하는데 필수 용어라 할 수 있다. 본 연구에서 제시한 이들 핵심 용어는 고등학교 수준에서의 과학이나 생물을 이해하는데 충분하며 또한 반드시 이해해야할 필수 항목일 것이다.

#### IV. 결론

고등학교 과학, 생물1, 및 생물2 교과서에 서술된 과학용어 네트워크를 구축하여 구조적 특성을 분석하였다. 과학교과서 용어의 빈도수 분포를 살펴보면, 물, 몸, 빛, 및 에너지와 같이 반복적으로 많이 사용되는 용어가 소수 있으며 사용되는 횟수가 극소수인 용어가 많이 있는 전형적인 척도 없는 네트워크 특성을 보여주었다 [그림 2].

복잡계 네트워크의 구조적 특성을 통해서 얻을 수 있는 정보는 한계가 있다. 복잡한 네트워크를 가중치가 낮은 링크를 점진적으로 제거하여 중요한 핵심 네트워크를 만든다면, 복잡계 네트워크에서는 발견되지 않는 유용한 정보를 얻을 수 있을 것으로 가정하고 이를 위해 k-core 알고리즘을 도입하였다[14].

K-core 알고리즘을 사용하여 복잡한 네트워크를 핵심 네트워크로 전환하여 핵심 과학 용어 네트워크를 구축하였다. 구축한 과학, 생물1, 및 생물2로부터 핵심 과학 용어를 추출하였는데, 이는 연결수가 많은 것부터 연결수가 낮은 것까지 일렬로 나열하여 일정한 링크수 이상을 가지는 용어를 추출하는 것과는 다르다. 왜냐하면 k-core 알고리즘을 사용하여 가중치가 낮은 것을 제거하는 방법 내에는 링크수뿐만 아니라 다른 용어와의 상관관계를 중시하는 멍침계수, 모듈을 형성하는 인자, 그리고 모듈과 모듈을 연결하는 인자까지 포함하기 때문이다. 따라서 본 연구에서는 구조 네트워크로부터 정보를 얻을 수 있는 정보 네트워크를 형상화하여 유용한 정보를 얻고자 하였다. 이러한 시도는 앞으로 복잡한 네트워크에 숨겨진 정보를 얻을 수 있는 가능성을 제시

하고 있다

본 연구에서 과학, 생물1, 및 생물2 교과서 각 과목당 3 교과서를 선정하여 과학용어 핵심 네트워크 [그림 3-5]를 만들었는데, 명확한 모듈 구조를 보여주는 교과서가 있는 반면에 그렇지 못한 교과서도 있었다. 모듈 구조가 명확하게 나타나는 교과서는 과목이나 단원이 비교적 잘 세분화되었다고 볼 수 있다. 일부 교과서 중에서 모듈을 형성하지 않은 것이 체계적으로 쓰이지 않았다는 것을 의미하는 것이 아니다. 이런 경우 과목과 과목 또는 단원과 단원사이에서 겹쳐지는 과학용어를 비중 있게 다루게 되면 그 또한 핵심 네트워크에서 모듈 구조를 나타내지 않을 수 있기 때문이다.

핵심 네트워크로부터 추출한 핵심 과학 용어들이 과학이나 생물의 개념을 정립하고 교과목을 이해하는데 필수 용어인 것은 사실이다[부록]. 따라서 고등학생들이 본 연구 부록에서 제시한 핵심 용어를 충분히 이해한다면, 고등학교 수준에서의 과학과 생물을 이해하는데 충분하다고 본다. 아울러 본 연구에서 제시한 이들 용어들을 데이터베이스화 한다면 과학 교과서 및 생물 교과서 네트워크를 바탕으로 상관관계를 분석하여 다양한 학습 분야에 모델로 활용할 수 있을 것으로 기대한다.

#### 참 고 문 헌

- [1] 안영은, "한국과 미국의 고등학교 생물교과서 비교분석-물질대사 단원을 중심으로-", 경희대학교 교육대학원 생물교육전공 석사학위논문, 2007.
- [2] R. A. Hill and R. I. M. Dunbar, "Social network size in humans," *Human Nature*, Vol.14, No.1, pp.53-72, 2003.
- [3] M. E. J. Newman and J. Park, "Why social networks are different from other types of networks," *Phys. Rev. E*, Vol.68, p.036122, 2003.
- [4] D. Liben-Nowell and J. Kleinberg, "Tracing information flow on a global scale using internet chain-letter data," *Proc. Nat'l. Acad. Sci.*, Vol.105, No.12, pp.4633-4638.
- [5] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and 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, and Z. N. Oltvai, "Lethality and centrality in protein networks," *Nature*, Vol.411, No.6833, pp.41-42, 2001.
- [7] 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.
- [8] Y. K. Lee, M. H. Yeo, T. H. Kang, J. S. Kim, and H. Y. Kim, "Identification of disease-associated proteins from atopy-related disease network," *J. Kor. Conten. Assoc.* Vol.9, No.4, pp.114-120, 2009.
- [9] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: Diameter of the world-wide web," *Nature*, Vol.401, No.6749, pp.130-131, 1999.
- [10] J. Stiller, D. Nettle, and R. Dunbar, "The small world of Shakespeare's plays," *Human Nature*, Vol.14, No.4, pp.397-408.
- [11] S. Kim, "Complex network analysis in literature: Togi," *Sae Mulli*, Vol.50, No.4, pp.267-271, 2005.
- [12] Y. K. Lee, H. I. Shin, J. E. Ku, and H. Y. Kim, "Analysis of network dynamics from the Romance of the three kingdoms," *J. Kor. Conten. Assoc.* Vol.9, No.4, pp.364-371, 2009.
- [13] C. N. Yoon and S. K. Han, "Network structure of physics-knowledge networks," *Sae Mulli*, Vol.53, No.5, pp.436-442, 2006.
- [14] J. Ignacio, A. Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, "K-core decomposition of internet graphs: Hierarchies, self-similarity and



measurement biases," Networks Hetero. Media, Vol.3, No.2, pp.371-393, 2008.

- [15] C. N. Yoon, D. Y. Lee, and S. K. Han, "Skeleton network analysis of the weighted network of science," Sae Mulli, Vol.55, No.6, pp.463-469, 2007.

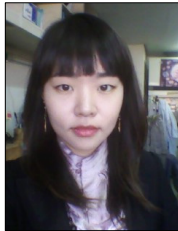
**저 자 소개**

**박 별 나(Byeol Na Park)** 준회원



- 2010년 2월 : 충북대학교 생화학  
과(이학사)
- 2010년 3월 ~ 현재 : 충북대학교  
생화학과 석사과정  
<관심분야> : 생물정보학, 신호전  
이 네트워크, 시스템생물학

**이 윤 경(Yoon Kyeong Lee)** 준회원



- 2008년 2월 : 충북대학교 생화학  
과(이학사)
- 2008년 3월 ~ 현재 : 충북대학교  
바이오정보기술학과 석사과정  
<관심분야> : 생물정보학, 신호전  
이 네트워크, 시스템생물학

**구 자 을(Ja Eul Ku)** 준회원



- 2009년 8월 : 충북대학교 생화학  
과(이학사)
- 2009년 9월 ~ 현재 : 충북대학교  
바이오정보기술학과 석사과정  
<관심분야> : 생물정보학, 신호전  
이 네트워크, 시스템생물학

**홍 영 수(Young Soo Hong)** 정회원



- 1980년 2월 : 고려대학교 식품공  
학과(농학사)
- 1982년 2월 : 고려대학교 농화학  
과(농학석사)
- 1993년 12월 : 오하이오주립대학  
교 생화학과(이학박사)
- 2009년 4월 ~ 현재 : 충북대학교 생화학과 BK교수  
<관심분야> : 분자생물학, 단백질 생화학, 단백질 네  
트워크

**김 학 용(Hak Yong Kim)** 종신회원



- 1985년 2월 : 충북대학교 농화학  
과(농학사)
- 1987년 2월 : 충북대학교 화학과  
(이학석사)
- 1994년 5월 : 미국 코네티컷대학  
교, 분자세포생물학과(이학박사)
- 1998년 3월 ~ 현재 : 충북대학교 생화학과 교수  
<관심분야> : 시스템생물학, 신호 전이, 단백질 네트  
워크, 생체동역학