

# 단백질 기능 흐름 모델 구성 및 평가 기법

## (A Method for Protein Functional Flow Configuration and Validation)

장우혁<sup>†</sup>      정석훈<sup>†</sup>

(Woo-Hyuk Jang)      (Suk-Hoon Jung)

한동수<sup>\*\*</sup>

(Dong-Soo Han)

**요약** 단백질 상호작용의 예측 및 실험 결과가 대용량으로 배포되면서 바이오 정보 기술 연구자들은 생명체 내의 단백질 상호작용 네트워크를 구성하기 위해 노력하여 왔다. 일반적으로 대용량의 상호작용 데이터들은 많은 오류를 포함한다고 알려져 있으나, 최근 단백질의 물리 화학적 특성 및 구조를 기반으로 한 방법들이 실제 실험과 병행되어 고품질(High resolution)의 결과를 제공하게 되면서, 특정 종에 대한 단백질 상호작용 네트워크가 점차 완성되고 있다. 그러나, 단순 물리적 링크 수준의 단백질 상호작용 네트워크만으로는 특정 병원체의 발병 매커니즘 규명 등과 같은 응용분야의 활용에 한계가 있다. 본 논문에서는 실험을 통하여 보고된 신호 전달 경로(signaling transduction pathway)를 이용하여 단백질 기능 간의 관계를 방향성이 있는 그래프로 표현한 단백질 기능 흐름 모델을 제시한다. 제안하는 모델은 Gene Ontology에서 정의된 molecular function을 정점(vertex)으로 가지고 이들 사이의 관계를 간선(edge)으로 표현함으로써 특정 기능의 전이를 살펴볼

수 있다. 이러한 기능 흐름 모델은 수 만개의 정점(vertex)으로 구성된 단백질 상호작용 네트워크에서 의미 있는 경로를 추출하는 데에 제약 혹은 참조 조건으로 사용될 수 있어 향후 활용도가 클 것으로 기대한다. 평가는 KEGG에서 제공되는 11개의 인간 신호 전달 경로 각각에 대하여 대상 경로를 제외한 나머지로부터 생성된 모델과의 크론바하 알파 계수(Cronbach's alpha)를 측정하였고( $\alpha=0.67$ ), 총 1023개의 흐름 중  $\alpha=0.6$  이상의 신뢰도에 대하여 총 765개의 흐름을 가지는 기능 흐름 모델을 최종 구성하였다.

**키워드** : 단백질 상호작용, 단백질 기능 흐름 모델, 신호 전달 경로, 크론바하 알파 계수

**Abstract** With explosively growing PPI databases, the computational approach for a prediction and configuration of PPI network has been a big stream in the bioinformatics area. Recent researches gradually consider physicochemical properties of proteins and support high resolution results with integration of experimental results. With regard to current research trend, it is very close future to complete a PPI network configuration of each organism. However, direct applying the PPI network to real field is complicated problem because PPI network is only a set of co-expressive proteins or gene products, and its network link means simple physical binding rather than in-depth knowledge of biological process. In this paper, we suggest a protein functional flow model which is a directed network based on a protein functions' relation of signaling transduction pathway. The vertex of the suggested model is a molecular function annotated by gene ontology, and the relations among the vertex are considered as edges. Thus, it is easy to trace a specific function's transition, and it can be a constraint to extract a meaningful sub-path from whole PPI network. To evaluate the model, 11 functional flow models of *Homo sapiens* were built from KEGG, and Cronbach's alpha values were measured ( $\alpha=0.67$ ). Among 1023 functional flows, 765 functional flows showed 0.6 or higher alpha values.

**Key words** : protein interaction, functional flow model, signal transduction pathway, Cronbach's alpha

## 1. 서론

계산적인 접근을 통한 단백질 상호작용 예측은 이를 직접 실험실에서 수행할 경우 요구되는 많은 시간과 비용을 줄일 수 있다는 장점을 가지고 있어 지난 몇 년간 다양한 방법이 제안되었다. 최근 단백질의 물리 화학적 특성 및 구조를 기반으로 한 방법들이 실제 실험과 병행되어 고품질의 결과를 제공하게 되면서, 조만간 특정 종에 대한 단백질 상호작용 네트워크의 전체 내역이 완성될 것으로 보인다. 그러나, 단순 물리적 링크 수준의

· 이 논문은 교육과학기술부, 한국산업기술재단, 한국과학재단의 특장기초 연구지원사업(학제기초) (No. R01-2008-000-20765-0) 및 지역혁신인력양성사업으로 수행된 연구결과임

· 이 논문은 제35회 추계학술대회에서 '단백질 기능 흐름 모델 구성 및 평가 기법'의 제목으로 발표된 논문을 확장한 것임

<sup>†</sup> 학생회원 : 한국정보통신대학교 공학부

torajim@icu.ac.kr

jsh@icu.ac.kr

<sup>\*\*</sup> 종신회원 : 한국정보통신대학교 공학부 교수

dshan@icu.ac.kr

논문접수 : 2008년 12월 19일

심사완료 : 2009년 2월 16일

Copyright©2009 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨팅의 실제 및 래터 제15권 제4호(2009.4)

단백질 상호작용 네트워크만으로는 특정 병원체의 발병 메커니즘 규명 등과 같은 응용분야의 활용에 한계가 있다. 단백질 상호작용 네트워크는 동시에 발견되는 단백질 혹은 유전 물질(gene product)의 집합으로, 네트워크의 링크는 단순한 물리적 결합(physical binding)만을 의미하며 생명 작용(biological process)의 상세 정보가 누락 되어 있기 때문이다. 단백질 상호작용 네트워크가 실제 응용분야에 활용되기 위해서는 단백질간의 활성, 비활성등과 같은 상호작용 링크의 상세가 추가되어야 한다. 뿐만 아니라, 의미 있는 대사 경로(예: 신호 전달 경로)가 어떠한 단백질에서 시작하여 어디에서 종결되는지를 알아내기가 쉽지 않으며, 그룹 혹은 단백질 복합체(protein complex)로서의 작용 여부 없이 단백질간의 일대일 대응만이 표현되어 있어, 추가정보가 필요한 실정이다.

Tong 그룹은 효모 단백질의 90%를 대상으로 synthetic genetic array(SGA)를 사용하여 전체 상호작용 네트워크를 구성하고, 각각 연결된 단백질 쌍의 특성 분석을 통하여 의미 있는 서브 네트워크를 추출하였다[1]. 이 연구에서는 전체 단백질 상호작용 쌍 중에 단지 12% 정도가 완전히 동일한 GO annotation을 공유했으며, 27%정도가 유사한 GO annotation을 공유한 것을 이용하여 상호작용 중에서 좀더 연관성이 강한 단백질 쌍을 구분하였다. Jacob Scott 그룹은 신호 전달 경로의 추출을 위해서 단백질 상호작용 네트워크상의 경로를 세포막(membrane)에서 nuclear 단백질, 그리고 최종적으로 transcription factor로 연결하는 방법을 제시하였다[2]. 이와 같은 연구들은 단백질 상호작용 네트워크에서 좀더 밀접한 관련이 있는 단백질 쌍을 찾아내는 기법 및, gene ontology를 이용한 정점(vertex) 구분 등은 단백질의 기능 분석이 의미 있는 서브 네트워크 추출을 위한 핵심 과점임을 나타내고 있다. 최근에는 상호작용 하는 단백질들이 가지는 기능을 중심으로 일정한 패턴을 찾고 [3], 단백질 기능 사이의 관계를 템플릿화 하여 대사 경로의 예측으로 확장하는 연구가 보고되었다[4]. 이와 같은 패턴 중심의 접근 방식은 적용의 용이성에도 불구하고 정보의 추상화로 인한 정보 손실과 패턴 내 기능 사이의 생물학적 의미 파악에 한계가 있다.

본 연구에서는 기존에 실험을 통하여 알려진 신호 전달 경로를 이용하여, 단백질 기능 사이의 관계를 방향성이 있는 그래프로 표현한 단백질 기능 흐름 모델을 제시한다. 제안하는 모델은 Gene Ontology에서 정의된 molecular function을 정점(vertex)으로 가지고 이들 사이의 관계를 간선(edge)으로 나타내고 있으며, 간선간에는 activation, inhibition과 같은 관계가 설정되어 있어 특정 기능의 전이를 파악하기 용이하다. 모델의 내부 일

관성 평가는 KEGG[6]에서 제공되는 11개의 인간 신호 전달 경로 각각에 대하여 대상 경로를 제외한 나머지로 부터 생성된 모델과의 크론바하 알파 계수(Cronbach's alpha)를 측정하였고( $\alpha=0.67$ ),  $\alpha=0.6$  이상의 신뢰도에 대하여 총 765개의 흐름을 가지는 기능 흐름 모델을 최종 구성하였다. 본 모델은 수 만개의 정점(vertex)으로 구성된 단백질 상호작용 네트워크에서 의미 있는 경로를 추출하는 데에 제약 혹은 참조 조건으로 사용될 수 있어 향후 활용도가 클 것으로 기대한다. 본 논문은 2장에서 기능 흐름 모델링의 방식 및 평가를 상세히 다루며, 실제 사용된 데이터에 대해서도 기술한다. 3장에서는 생성된 모델에 대한 특징과 보완점등을 논의하고 마지막 4장에서 향후 활용방안에 대해 기술한다.

## 2. 실험 방법 및 결과

### 2.1 기능 흐름 구성

단백질 기능 흐름 모델은 단백질 사이의 관계를 각 단백질에 해당하는 기능 사이의 관계로 확장 시킨 개념이다. 본 연구에서는 실험적으로 알려진 신호 전달 경로에서 단백질과 단백질 상호작용 관계 정보를 추출하였다. 신호 전달 경로는 외부의 자극에 대한 세포 내 기작을 나타낸 것으로 생명현상을 이해 할 수 있는 기초 자료이다. 대부분의 신호 전달 경로에서는 일반적인 대사 경로와 다르게 단백질과 이들의 관계로 구성되어 있어, 기능 흐름 모델을 작성하기 용이한 측면이 있다. 단백질에 해당하는 기능은 Gene Ontology의 세가지 카테고리 (cellular component, biological process, molecular function) 중 molecular function을 사용하였다. 그림 1은 Hedgehog 신호 전달 경로의 일부 단백질과 그 관계 등이 나타나 있다. 각각의 단백질은 일반적인 기능이 보고되어 있으며 이들 사이에는 inhibition, dissociation과 같은 관계가 형성되어 있다.

단백질 기능 흐름 모델은 신호 전달 경로와 같은 의미 있는 단백질 상호작용 경로에는 단백질 기능 사이에 특징적인 흐름이 존재한다는 것을 기본 가정으로 한다. 이러한 관점에서 특징적인 단백질 기능 흐름을 보이는 단백질 사이의 관계(예: activation)는 동일한 기능 흐름을 보이는 또 다른 단백질 쌍에도 동일하게 적용될 수

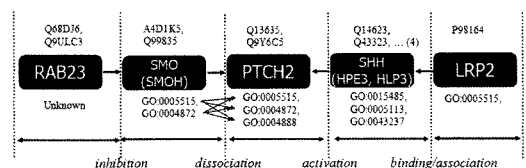


그림 1 Some proteins and their relationship in Hedgehog Signaling Pathway

있으며, 의미 있는 단백질 상호작용 경로에서 나타나는 기능의 흐름은 각각 이전 단계의 이웃 단백질 기능에 따라 선택적으로 결정된다. 예를 들어, LRP2 단백질이 SHH 단백질에 binding 한다는 신호전달 경로의 정보를 이용하여 LRP2의 기능인 protein binding (GO:0005515) 과 SHH의 기능인 cholesterol binding (GO:0015485), patched binding (GO:0005113), laminin-1 binding (GO:0043237) 들 사이는 모두 binding/association 관계를 가진다고 확장하는 것이다. 이러한 기능 사이의 관계가 하나의 단백질 쌍에서 중복되어 나타나는 경우 1회 출현으로 처리하고, 서로 다른 단백질 쌍에서 중복되는 경우는 출현 빈도를 누적하여 기능 흐름의 weight로 사용하였다. 기능이 알려지지 않은 단백질들은 모델 구성에서 제외하였다. 이와 같은 방식으로 KEGG에서 제공하는 인간 신호 전달 경로 11종의 단백질 상호작용에서 총 1023개의 기능 흐름을 추출하였다.

표 1에서와 같이 신호 전달 경로상의 한 정점은 n개의 KEGG ID와 UniProt ID[5]를 가질 수 있으며, UniProt ID 한 개당 다시 n개의 molecular function을 가질 수 있다. 기능 흐름 모델 작성 시, 신호 전달 경로의 한 정점에 해당하는 기능은 그 정점에 해당하는 모든 기능의 합집합으로 사용하였다. 기능 흐름 모델은 그림 2와 같이 핵심 기능을 중심으로 방사형으로 주변 기능이 연결되어 있는 형태를 띄고 있다.

2.2 검증

생성된 기능 흐름 모델의 검증은 크론바하 알파 계수 [7]를 이용하여 내적 일관성 분석을 수행하였다. 본래 크론바하 알파 계수는 한 개념을 많은 항목으로 측정했을 때, 각 항목들에 대한 일관성이나 동질성을 측정하는 것으로 항목들간의 상관관계를 통해 평가된다. 이때 상관관계가 높을수록 내적 일관성이 높고, 낮을수록 내적 일

표 1 Statistical values from KEGG signaling transduction pathways and functional flows

ID	tC	tR	aK	aU	aF	aFC	tFR	aFR
04010	136	171	2.17	2.85	3.55	25	255	10.2
04012	80	86	1.5	1.85	2.78	12	90	7.5
04310	83	70	2.34	2.94	2.76	13	118	9.08
04330	31	16	1.77	2.19	2.19	3	28	9.33
04340	25	16	2.64	3.16	2.72	5	29	5.8
04350	70	45	1.8	2.43	3.49	9	85	9.44
04370	40	34	2.13	2.7	3.2	8	89	11.13
04630	33	27	4.85	6.58	6.15	7	154	22.0
04020	62	31	3.31	4.21	4.06	15	51	3.4
04070	66	77	2.38	2.77	2.88	12	186	15.5
04150	44	31	1.55	1.91	2.20	7	52	7.43

- ID: 04010 MAPK 04012 ErbB 04310 Wnt
- 04330 Notch 04340 Hedgehog 04350 TGF-beta
- 04370 VEGF 04630 Jak-STAT 04020 Calcium
- 04070 Phosphatidylinositol 04150 mTOR
- tC : 신호 전달 경로 상의 총 정점 수
- tR : 신호 전달 경로 상의 총 간선 수
- aK : 정점 당 KEGG ID의 평균 개수
- aU : 정점 당 UniProt ID의 평균 개수
- aF : 정점 당 Gene Ontology의 molecular function 개수
- aFC : 기능 흐름 모델의 총 정점 수
- tFR : 기능 흐름 모델의 총 간선 수
- aFR : 기능 흐름 모델의 정점당 간선 평균

관성이 낮다고 평가하게 되며, 계산 공식은 다음과 같다.

$$\alpha = \frac{N}{N-1} \left( 1 - \frac{\sum_{i=1}^N \sigma_{Y_i}^2}{\sigma_X^2} \right)$$

단, N=문항수,  $\sigma_{Y_i}^2$  = 각 문항의 분산,  $\sigma_X^2$  = 총 분산  
 크론바하 알파 모델의 각 변수는 단백질 기능 흐름 모델에서 다음과 같이 대응된다.

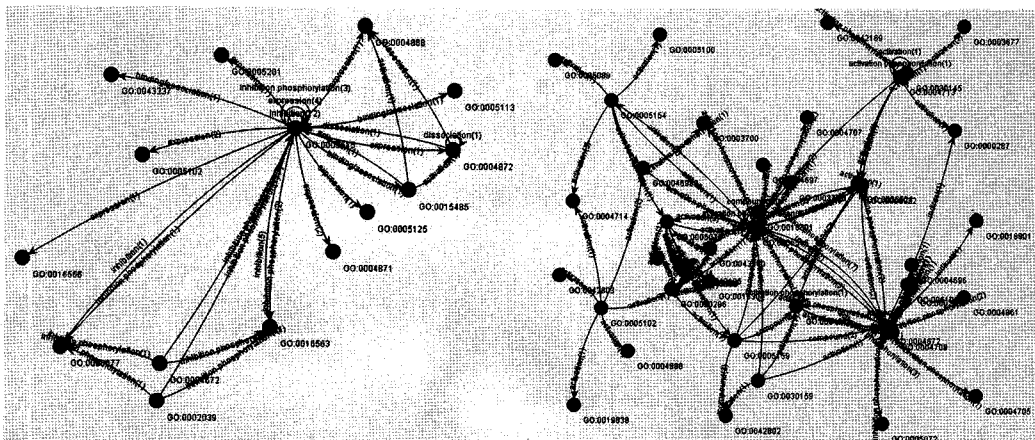


그림 2 Functional flow diagrams of Hedgehog and ErbB signaling pathways

- $N$  = 해당 신호 전달 경로에서 추출된 단백질 기능 흐름의 총 수
- $\sigma_{f_i}^2$  = 특정 단백질 기능 흐름의 총 11개 신호 전달 경로에 대한 분산
- $\sigma_x^2$  = 해당 신호 전달 경로내의 총 단백질 기능 흐름의 분산

이때, 해당 신호 전달 경로의 특정 단백질 기능 흐름이 다른 10개의 신호 전달 경로에서 다른 타입으로 발견될 경우 기능 흐름의 출현 빈도수의 합은 최대값 11에서 감소시키는 방법을 취하였다. 이 부분은 향후 개선의 여지가 있으나, 본 연구에서는 특정 단백질 기능 흐름이 다른 신호 전달 경로에서 나타나지 않는 경우에도 그 흐름은 자체로 의미가 있다고 간주 하였고, 서브 타입(예: activation) 이 다른 경우에만 충돌이 일어나는 것으로 처리하였다. 각 신호 전달 경로에서 추출된 단백질 기능 흐름의 크론바하 알파 계수 평균은 다음과 같으며, 신뢰도 계수 0.6 이상으로 구성된 기능 흐름의 총 개수는 표 2와 같다. 검증 결과 전체 단백질 기능 흐름의 크론바하 알파 계수 평균은 0.67로 비교적 신뢰할 수 있는 정도임을 알 수 있었으며, 알파 계수 0.6 이상의 중복을 제외한 단백질 기능 경로는 총 765개가 발견되었다.

표 2 Validation result for each functional flow

ID	tFR	평균빈도	평균분산	크론바하	a>0.6
04010	255	9.43	7.56	0.63	211
04012	90	9.76	6.03	0.72	78
04310	118	8.41	11.29	0.48	36
04330	28	8.68	8.50	0.60	17
04340	29	8.31	14.15	0.54	10
04350	85	9.27	8.46	0.66	73
04370	89	8.96	6.31	0.64	50
04630	154	9.33	6.03	0.65	75
04020	51	10.24	7.24	0.86	47
04070	186	10.76	1.07	0.87	169
04150	52	9.56	7.05	0.71	46

2.3 데이터베이스

신호 전달 경로의 단백질들과 연관 관계를 얻기 위하여 KEGG의 SOAP/WSDL 인터페이스(released on August 17, 2007)를 사용하였고, 이 인터페이스를 통해 신호 전달 경로 Release 47.0, July, 2008의 데이터를 사용하였다. 단백질 아이디는 범용적으로 사용되는 UniProt ID(release 14.0, July 22, 2008)를 사용하였다. 마지막으로 Gene ontology에서 제공하는 lite version의 term을 molecular function으로 사용하였다(release on July 1, 2008).

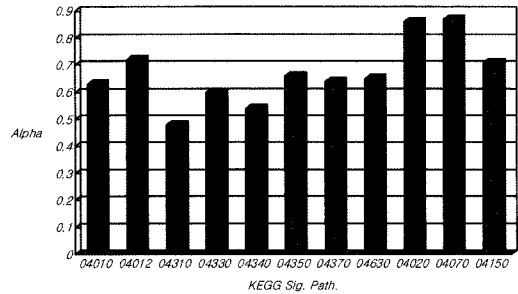


그림 3 Cronbach's alpha values for each functional flow extracted from signaling transduction pathways

3. 고찰

단백질 상호작용 네트워크에서 신호 전달 경로와 같은 의미 있는 서브 네트워크를 찾는 연구는 주로 주어진 하나의 노드에서 종결 노드로의 simple path를 찾는 문제로 간주된다[1,2]. 이에 기존 연구들은 하나의 단백질 쌍과 연결된 이웃과의 관계를 주로 알려진 문헌 정보의 개수 혹은 Gene Ontology의 유사성 등을 이용하여 확률 값으로 표현하였다. 이러한 방식은 각각의 단백질 쌍 사이에서 측정된 유사성이 다음 경로와 연동되지 않는다는 점에서 일정한 흐름을 가지지 못하였다. 본 연구에서는 단백질 기능 사이의 일정한 흐름을 미리 모델링 하여, 단백질 상호작용 네트워크의 경로 탐색에 제약 조건으로 사용할 수 있는 기능 흐름 모델을 제안하였다. Ali Cakmak[4]은 Pathway Functionality Template(PFT)라는 모델을 이용하여 특정 기능의 연결을 패턴으로 찾아내는 시도를 하였다. PFT는 본 논문에서 제안하는 기능 흐름 모델과 상당 부분 유사하나, 몇 가지 다른 측면이 있다. PFT는 Gene Ontology가 Hierarchical한 특성을 지님을 이용하여, 패턴 상의 기능들을 최대한 조상 노드의 (좀 더 일반적인) GO term으로 치환하고 단순화하였다. 반면, 제안하는 기능 흐름 모델은 가장 하위 노드 (좀 더 자세한) GO term을 그대로 사용하였으며, 특정 패턴을 찾기 보다 신뢰할 수 있는 범용적인 모델을 작성하고자 하였다. 이 부분에서, 알려진 실험에서의 출현 빈도가 매우 낮더라도 이를 추상화하기 보다 그대로 차용하고자 하였다. 이러한 이유로 생성된 모델은 표 1에서와 같이 tFR의 총합은 1137인데 비해 중복되지 않는 기능 흐름은 1023으로 신호 전달 경로 사이에 중복이 그다지 크지 않음을 보여준다. 이는 단백질의 기능으로 사용한 molecular function 의 추상화 레벨의 조정으로 변경할 수 있다. 예를 들어 Mehmet E Turanalp 의 연구[3]에서는 slim GO를 사용하여 총 기능 개수를 50개 미만으로 조정하여 사용하

였다. 생성된 기능 흐름 모델의 생물학적 유효성 검증은 실제 단백질 상호작용 네트워크의 경로와 기능 흐름 모델의 교차 비교를 통하여 수행할 수 있으나 향후 과제로 남겨둔다. 검증에서 사용한 크론바하 알파 계수는 내적 일관성을 측정하는 대표적인 기법이다. 이 방법에서는 하나의 개념을 가진 다양한 질문의 일관성을 측정하며, 질문들 사이의 중요도는 동일하다고 간주한다. 사실 단백질 기능 흐름들은 출현 빈도 수로 중요도를 구분할 수 있으나, 출현 빈도가 낮은 것과 높은 것을 생물학적 중요도의 차이가 있다고 할 수는 없다. 따라서 검증을 위해 각각의 단백질 기능 흐름은 동일한 정도의 중요도를 가진다고 가정하였고, 이를 바탕으로 평균 0.67의 신뢰도가 측정되었다. 크론바하 알파 계수 측정법에서는 0.8-0.9의 경우 상당히 높은 신뢰도, 0.7-0.8의 경우 적당한 신뢰도, 0.6-0.7은 수용 가능한 신뢰도로 구분하고 있다. 대용량의 단백질 관련 데이터베이스의 현재 오류 포함 정도를 감안할 때, 본 연구의 검증 결과는 매우 긍정적인 결과라고 할 수 있다.

한편, 그림 3에서와 같이 각각의 신호 전달 경로는 서로 다른 크론바하 알파 계수를 보여 준다. 여기서 크론바하 알파 계수는 해당 신호 전달 경로에 존재하는 기능 흐름들의 형태가 다른 신호 전달 경로에서도 유사하게 나타남을 의미하며, 서로 배치되는 기능 흐름이 적게 발견 될수록 크론바하 알파 계수는 높은 값을 보여준다. 본 논문에서 각각의 신호 전달 경로가 보여주는 크론바하 알파 계수는 차이가 있음을 밝혔는데, 이는 기능 흐름 모델의 오류라기 보다 gene ontology가 가지는 추상화 레벨의 차이에서 기인한다. GO:0005515(protein binding)와 같은 추상적인 기능은 연구가 성숙하지 못하여 좀더 세밀한 기능이 밝혀지지 않았기 때문에, 일부 기능 흐름의 충돌이 일어날 수 있다. 이를 증명하기 위하여 각각의 신호 전달 경로에서 작성한 기능 흐름 모델의 gene ontology 추상화 레벨을 측정하였다. 추상화 레벨은 기능 흐름 모델에 해당하는 기능들의 gene ontology hierarchy 상 root 노드로부터의 평균 거리로 계산하였다. 그 결과 그림 4에서와 같이 전반적으로 크론바하 알파 계수와 평균 거리의 분포 경향은 유사한 경향성을 보였다. 이는 반대로 단백질 기능 흐름 모델의 작성에 있어서, 신호 전달 경로상에 나타나는 단백질 기능의 추상화 레벨의 측정은 기능 흐름 모델의 신뢰도를 가능하게 할 수 있는 척도가 될 수 있음을 암시하고 있다.

#### 4. 결론 및 향후 과제

본 연구에서는 실험을 통하여 보고된 신호 전달 경로로부터 단백질 기능 중심의 기능 흐름 모델을 구성하였고, 크론바하 알파 계수를 이용하여 내적 일관성을 검증

GO term distance from root

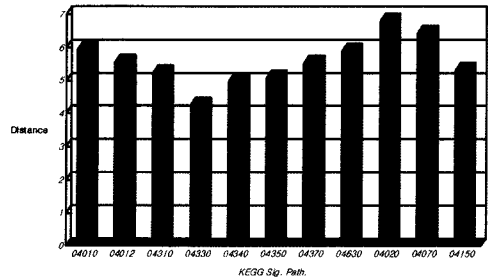


그림 4 Average distances of GO terms from root node on GO hierarchy

하였다. 검증 결과 구성된 기능 흐름 모델은 알파 계수 0.67로 수용 가능한 정도의 신뢰도를 보였고, 전체 기능 흐름 1023개 중 알파 계수 0.6 이상은 765개였다. 본 연구는 복잡한 단백질 상호작용 네트워크에서 특정 기능의 흐름을 중심으로 한 의미 있는 서브네트워크 추출 등에 중요한 참조 모델로 사용될 수 있을 것으로 기대한다. 또한 향후 신호 전달 경로와 같은 대사 경로 예측 연구 분야에 중요한 선행 연구로서의 가치를 가진다. 본 연구진은 향후 제안한 모델을 이용하여, 기존에 알려진 신호전달 경로의 보완 및 새로운 신호전달 경로의 예측, 나아가 대사 경로의 예측 등을 수행할 예정이다.

#### 참고 문헌

- [1] Amy Hin Yan Tong, "Global Mapping of the Yeast Genetic Interaction Network," *Science*, 303:5659, 2004.
- [2] Jacob Scott, et. al., "Efficient algorithms for detecting signaling pathways in protein interaction networks," *Journal of Computational Biology*, 13:2, 2006.
- [3] Mehmet E Turanalp and Tolga Can, "Discovering functional interaction patterns in protein-protein interaction networks," *BMC Bioinformatics*, 9:276, 2008.
- [4] Ali Cakmak and Gultekin Ozsoyoglu, "Mining biological networks for unknown pathways," *BIOINFORMATICS*, 23:20, 2007.
- [5] The UniProt Consortium, "The Universal Protein Resource (UniProt)," *Nucleic Acids Res.*, 36:D190-D195, 2008.
- [6] Kanehisa, M., et. al., "KEGG for linking genomes to life and the environment," *Nucleic Acids Res.*, 36:D480-D484, 2008.
- [7] Lee J. Cronbach, et. al., "The dependability of behavioral measurements," *Science*, 178:4067, 1972.