

MAPK Hypotonic Shock의 Signaling Pathway에 대한 시뮬레이션

조미경*, 서정만**, 박현석***

Simulation for Signaling Pathway of MAPK Hypotonic Shock

Jo Mi Kyung*, Seo Jeong Man**, Park Hyun Seok***

요약

Yeast를 이용하여 Two-Hybrid System 실험을 통해 밝혀진 단백질 상호작용 데이터에 단백질 위치 정보를 이용하여 가중치를 부여하고 단백질 신호 전달 경로를 추출하였다. 그 결과 중 MAPK Hypotonic Shock 기능의 데이터를 가지고 KEGG에서 제공하는 신호전달 경로와 비교하여 어느 정도 일치 하는지의 유사도를 측정하고 시뮬레이션 하였다. 이때 프로세스 실행 시간도 측정하여 제시 하였다. 향후 연구를 발전시키면 다양한 유전적 질병의 원인과 치료제 개발의 단서를 제공할 수 도 있으며 더 나아가 신약 개발을 할 수 있다.

Abstract

We extracted protein signal delivery path from protein interaction data, using location information and weight of protein. We obtained the protein interaction data by experimenting in two-hybrid system using Yeast. We simulated function's data of Hypotonic Shock comparing to signal delivery path provided in KEGG from the results. We measured process running period as well. In future, this research can be key to discover the origin of various genetic diseases and develop treatment.

▶ Keyword : 상호작용(PPI), 위치정보(Localization), 신호전달(Signaling Pathway), 경로예측(Predicting), 가중치(Weight)

• 제1저자 : 조미경 교신저자 : 박현석

• 투고일 : 2009. 05. 23, 심사일 : 2009. 05. 26, 게재확정일 : 2009. 05. 28.

* 이화여자대학교 박사 ** 한국재활복지대학 교수 *** 이화여자대학교 교수

I. 서론

세포의 생명현상은 단백질 간 상호작용을 통해서 일어난다. 효모와 초파리 등과 같은 대표적인 실험 모델에 있어서는 그 종이 가지고 있는 모든 단백질(n)에 대한 상호작용 여부를 실험하는 종 수준 데이터(genome scale)들이 출현하고 있다 [1][2][3].

종 수준 단백질 상호작용 데이터는 단백질을 노드로 상호작용을 에지로 표현하였을 때, 소수의 연결도 높은 노드들이 존재하는 Scale Free Network[4]의 특징을 지니며 이러한 연결도 높은 노드들의 존재는 임의의 단백질 간 거리를 줄이는 역할을 한다.

또한 단백질 상호작용 네트워크는 Budding Yeast (Saccharomyces Cerevisiae)의 경우 제일 큰 클러스터 네트워크에 전체 단백질의 약 78%를 포함하는 것으로 알려져 있다[5].

신호 전달 경로란 세포가 외부 자극에 반응하여 새로운 단백질을 만들어 내야 할 필요가 있을 때 그 신호를 세포 표면으로부터 단백질 주형에 해당하는 DNA가 존재하는 세포핵 안쪽으로 전달하는 일련의 과정을 말한다. 이러한 신호 전달 과정은 필연적으로 단백질 간의 상호작용을 기반으로 한다. 이러한 신호 전달 과정을 통하여 외부 자극은 세포 전체로 그 자극을 확대하고 퍼뜨리는 것이 가능하다[6]. 신호 전달 경로는 전체 단백질 상호작용 네트워크의 부분 그래프에 해당한다.

세포 내 단백질 위치는 단백질의 기능을 유추할 수 있는 중요 정보의 하나로 인식되고 있다. 진핵생물의 경우 막으로 둘러싸인 세포 내 소기관들을 가지고 있으며 구분된 기관들은 서로 다른 기능을 수행하도록 분화되어 있다. 따라서 특정 단백질의 세포 내 위치를 아는 것은 단백질 연구에 있어 기본 정보가 된다.

세포 내 단백질 위치는 물리화학적 경계가 되기도 한다. 이는 서로 만날 필요가 없는 단백질 간에 불합리한 접촉을 피할 수 있어 효율성을 높이는 한 방법이 된다.

이처럼 생물학적 여러 의미 관계 구조들을 전산학적 관점에서는 네트워크로 접근이 가능하며 이론을 배경화한 주제로 다양한 그래프 이론들이 적용되고 있다. [4]

논문에서 다루고자 하는, MAP(Mitogen-Activated Protein) Kinase 신호전달경로는 세포막 단백질에 세포분열 유도물질인 Mitogen이 결합하며 시작된다. 이러한 신호전달 결과 세포 분화, 분열, 생존, 사망 등의 현상이 일어난다. 세포막에서 신호를 전달받은 단백질들은 두 개가 하나의 단위로 작용하는 이합체(Dimer)가 됨으로써 활성화부위가 노출된다. Tyrosine Kinase로서 활성화되어 단백질들을 인산화 하는 일련의 반

응을 유도한다. 차례로 인산화 되는 신호화 과정을 전달받은 MAP Kinase는 세포질에서 핵으로 이동하며 세포가 필요로 하는 새로운 단백질을 만들 수 있도록 한다. 이러한 MAP Kinase 신호전달 경로는 세포막에서 발생한 신호를 핵 안쪽까지 증폭하면서 전달하고, 각 단계에서 어느 단백질을 활성화 하느냐에 따라 다양한 반응을 나타낼 수 있다. 본 연구에서는 Yeast 단백질간 상호작용 정보와 세포 내 위치 정보를 활용하여 세포막 단백질로부터 핵 단백질까지 신호 경위를 찾아 Hypotonic Shock의 Signaling Pathway에 대한 시뮬레이션을 하고자 한다.

II. 신호전달 경로 추출 관련 연구

그래프 클러스터링 알고리즘(Graph Clustering Algorithm)을 적용하여 단백질 Complex와 Cellular Process에 해당하는 기능적 모듈을 찾은 선행연구가 존재한다[7]. 전체 그래프에서 다른 부분에 비해 연결도가 높은 지역[8]을 찾아 그 결과가 앞의 이론과 비슷하게 단백질 복합체거나 동일한 생물학적 과정에 참여하는 기능적 단위였음을 밝힌 연구도 있었다.

선행연구에서, 저자들은 다른 어느 정보도 사용하지 않고 그래프 알고리즘만을 적용하여 생물학적으로 의미 있는 부분을 추출하였다.

신호 전달 경로가 세포 표면에서 시작되어 핵 안으로 전달되는 이론에서 착안하여 시작점과 도착점을 정해놓고 상호작용 데이터와 단백질 발현 데이터를 사용하여 신호전달 경로를 추출하였다.

한편, 단백질 상호작용을 이용하여 경로를 찾으려는 시도가 있었으나 실험결과가 정확하게 나오지 못한 연구도 있었다[9].

DIP (Database of Interacting Proteins) [10]은 Yeast Two-Hybrid 혹은 면역침전 등 실험적 방법으로부터 얻은 단백질간의 상호작용 데이터를 제공한다. DIP Database는 일관성 단백질 상호작용 데이터베이스를 만들기 위해 Swiss-PROT과 다양한 리소스로부터 해당 단백질들에 대한 정보를 종합한다.

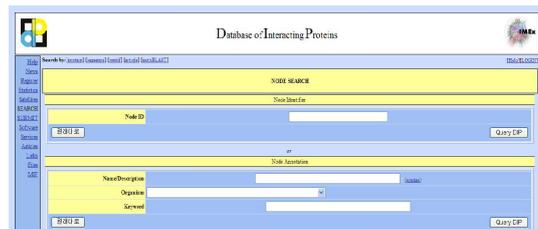


그림 1. DIP 데이터베이스
Fig 1. DIP Database

MIPS(Munich Information Center of Sequences)(11)의 CYGD는 *saccharomyces cerevisiae*의 Genomic Structure에 대한 정보를 제공하는데, 최근 Report, Graphical Displays, Genome의 특정 부분에 대한 Summary Table 등의 정보를 얻을 수 있다.

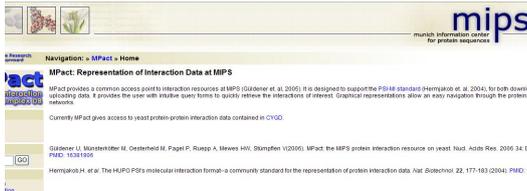


그림 2. MIPS 데이터베이스
Fig 2. MIPS Database

YPLD(Yeast Protein Localization Database) (12)는 Yeast(*saccharomyces cerevisiae*) 단백질의 Subcellular Localization에 관련된 정보를 제공한다. SubCellular Localization 정보는 GFP로 표시된 단백질을 편광 현미경 혹은 공초점 현미경을 통해 얻어진다.

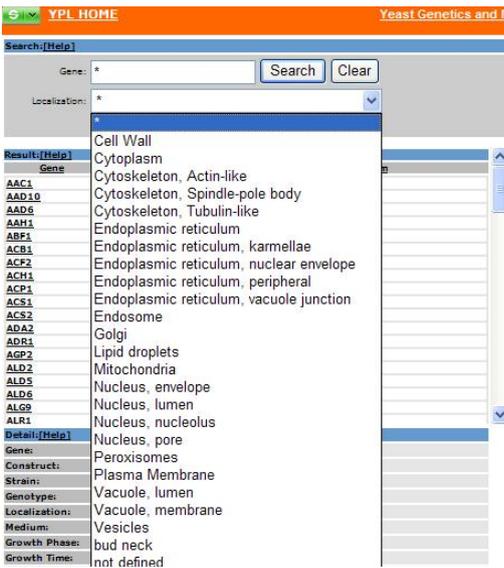
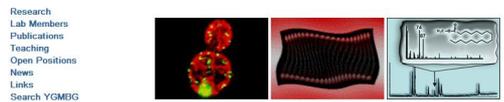


그림 3. YPLD 데이터베이스
Fig 3. YPLD Database

연구는 세포 내 단백질 위치가 알려진 Budding Yeast를 대상으로 하여 단백질이 여러 위치에 존재할 수 있다는 위치 정보를 활용한다. 이때 임의의 두 단백질 간 다양하게 작용하는 상호작용 빈도수와 단백질 위치의 단백질 수 확률 값을 적용하여 가중치로 부여 하였다.

가중치가 부여된 값 중 의미 있는 높은 값을 채택하는 방식으로 신호전달 경로를 추출한다. 최종적으로 KEGG(13)에서와 같이 기존에 실험을 통해 이미 알려진 신호전달 경로와 같은 경로를 추출 하는지 유사도 비교를 하였다. 단백질의 위치 정보는 다음 (그림 4)와 같다.

	A	B	C	D
1	YAL001C	cytoplasm	nucleus	
2	YAL002W	endosome		
3	YAL005C	cytoplasm		
4	YAL007C	ER		
5	YAL029C	ambiguous	bud neck	cytoplasm

그림 4. 단백질별 위치 정보
Fig 4. Information about Protein Position

단백질 상호작용 데이터는 DIP에서 제공하는 34,797개와 MIPS에서 제공하는 30,429개를 통합한다. 그 중 중복 데이터 19,153개를 제거하여 고유 상호작용 단백질 46,073개와 논문을 통해 밝혀진 의미 있는 데이터 48개를 추가하여 총 46,121개의 단백질 상호작용 데이터를 입력 데이터로 사용 하였다.

고유한 단백질 약 5,000여 개가 21개의 단백질 위치(14) 영역에 분포한다. 단백질 하나가 여러 위치에 존재할 수 있는 다중 분포 자료 7,787개의 위치정보 데이터를 사용 하였다.

앞서 언급한 바와 같이 경로 추출 단백질 모델로는 Yeast에서 Hypotonic Shock 기능과 연관되는 MAPK (Mitogen-Activated Protein Kinases) 신호 전달 경로를 대상으로 설정 하였다.

III. 세포 내 위치정보 활용 알고리즘

3.1 단백질 간 상호작용 경로 찾기

프로그램 실행 화면에서 시작 단백질(예:YLR332W)과 종료 단백질(예:YPL089C) 그리고 깊이(예:8)를 실행화면의 입력 값으로 준다. 단백질 간 상호작용 데이터 46,121개를 입력 자료로 하여 DFS(Depth First Search) 알고리즘을 수행하며 이때 입력받은 시작 단백질로부터 출발하여 제한 깊

이를 비교해 가며 핵에 존재하는 목표 단백질까지 모든 신호 전달 가능 경로를 찾게 된다. 예를 들어 탐색경로를 살펴보면 YLR332W를 출발 단백질로 하여 YOL109W-YBR069C-YPR198W-YHR026W-YJR091C-YDR167W를 찾고 마지막 목표 단백질인 YPL089C의 순으로 신호전달 경로를 추출 한다.

	A	B	C	D	E	F	G	H
1	start : YLR332W, target : YPL089C, maximum length : 8							
2	YLR332W	YOL109W	YBR069C	YPR198W	YHR026W	YJR091C	YDR167W	YPL089C
3	YLR332W	YOL109W	YBR069C	YLR372W	YBR054W	YJR091C	YDR167W	YPL089C
4	YLR332W	YOL109W	YBR069C	YLR372W	YJR010C	YJR091C	YDR167W	YPL089C
5	YLR332W	YOL109W	YBR069C	YLR372W	YML123C	YJR091C	YDR167W	YPL089C
6	YLR332W	YOL109W	YBR069C	YLR372W	YHR026W	YJR091C	YDR167W	YPL089C
7	YLR332W	YOL109W	YBR069C	YLR372W	YMR058W	YDL029W	YHR030C	YPL089C
8	...							
9	(종료)							

그림 5. 단백질 간 상호작용 결과
Fig 5. Result of Interactions between Proteins

3.2 가중치 적용

단백질 위치 정보의 중요성을 인식하고 위치 정보를 활용 하였다. “단백질 간 상호작용 경로 찾기” 단계에서 산출된 결과를 기반으로 모든 경로에 위치별 가중치와 상호작용 가중치를 이용한 평가함수를 적용하여 평가한다. “단백질 간 상호작용 경로 찾기”를 평가하기 위해서는 패스를 이루는 단백질 앞 뒤 순서 간 모든 평가 값을 곱한다. 단백질의 위치정보와 위치 간 확률 값을 담은 이차원 배열 정보를 사용 하였다. 이때 단백질 위치 가중치 계산은 21개의 위치에 따른 각 위치별 단백질 개수의 확률 값을 사용 한다. 단백질 간 상호작용 확률 값은 배열 (21위치X21위치)을 이용하여 두 단백질이 갖는 모든 위치의 상호작용 경우를 계산한 확률 값을 사용 하였다. 예를 들어 단백질의 위치정보가 YMR088C : vacuolar membrane, cytoplasm YLL040C: mitochondrion, endosome, cytoplasm 일때 YMR088C와 YLL040C의 모든 상호작용 경우의 수는 위치 개수 2*3한 값인 6개가 발생하는 것이다. 산출물은 (그림 6)에서 제시 한다.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Y56A3A2	YGL178W										
2												
3	YAL001C	YBR123C										
4	nucleus(1)-nucleus(1)[mitochondrion(20)-nucleus(1)]cytoplasm(5)-nucleus(1)											
5	YAL002W	YKR026C										
6	endosome(13)-cytoplasm(5)											
7	YAL002W	YLR148W										
8	endosome(13)-golgi(10)[endosome(13)-vacuole(19)]endosome(13)-vacuolar membrane(16)[endosome(13)-endosome(13)]											
9	YAL002W	YLR396C										
10	endosome(13)-vacuole(19)[endosome(13)-vacuolar membrane(16)]endosome(13)-cytoplasm(5)											
11	YAL002W	YMR231W										
12	endosome(13)-vacuole(19)[endosome(13)-vacuolar membrane(16)]endosome(13)-endosome(13)											
13	YAL002W	YPL049W										
14	endosome(13)-vacuole(19)[endosome(13)-vacuolar membrane(16)]											
15	YAL003W	YKL081W										
16	cytoplasm(5)-mitochondrion(20)[cytoplasm(5)-cytoplasm(5)]											

그림 6. 위치 인터랙션의 모든 경우의 수
Fig 6. All case of Position Interaction

단백질 위치정보와 단백질 상호작용 값을 이용하여 평가 함수인 두 가지 형태의 함수를 정립하였다. 첫째는 A-B 위치 상호작용 확률 값 둘째는 A위치의 가중치 * B위치의 가중치 * A-B 위치 상호작용 확률 값이다. 단백질I과 단백질J의 상호작용 가중치 확률 값(1)은

$$(a,b) \text{ Protein } i \text{의 위치} \times \text{Protein } j \text{의 위치} \quad (\text{Local a와 Local b의 Interaction 확률})$$

에 의해서 구하며 아래에서 방법1이라 칭한다. 단백질I와 단백질J의 상호작용 가중치 확률 값(2)는

$$(a,b) \text{ Protein } i \text{의 위치} \times \text{Protein } j \text{의 위치} \quad (\text{Local a와 Local b의 Interaction 확률} \times \text{Local a의 위치 가중치} \times \text{Local b의 위치 가중치})$$

에 의해 산출 하며 아래에서 방법2라 칭한다. 단백질1과 단백질2의 단백질 간 상호작용 값을 평가하기 위해서 단백질1과 단백질2의 단백질 위치 정보를 가져와 산출한다.

첫째 만약 단백질1의 위치가 없거나 단백질2의 위치가 없을 경우 1.0의 값을 평가한다. 둘째 단백질1과 단백질2 모두 하나 이상의 위치를 가질 경우 그 위치 간 쌍의 모든 경우의 수 중에서 평가함수 정보를 통해 산출한 값이 가장 큰 값을 선택 한다. 이렇게 선택된 단백질 간 상호작용 값들을 곱함으로써 한 개의 “단백질 간 상호작용 경로 찾기” 단계를 평가하게 된다.

3.3 통계 및 분석

가중치를 적용하여 구한 단백질 상호작용 경로 값을 내림차순으로 정렬하여 동일한 순위 내에 여러 개의 값이 존재할 경우 그 동일 순위의 값 중 최대값의 포함노드를 포함하고 있는 값을 추출하여 KEGG신호전달 경로와의 유사도를 비교분석 한다.

YLR332W-YPR165W-YBL105C-YJL095W-YOR231W-YHR030C-YER111C인 단백질 신호전달 경로 추출을 목표로 하여 시작 단백질을 YLR332W로 목표 단백질을 YER111C로 하고 깊이를 7로 하는 프로그램을 실행한 후 가중치 알고리즘을 적용하였다. 이때 산출된 결과 값을 내림차순 정렬하여 중복 값을 제거한 후 유일한 값에 대한 순위를 부여하고 최대값을 선출한다. 이때 최대값에 해당하는 것이 여러 개의 중복 값이 산출된 원인으로는 단백질이 여러 위치 정보를 가지고 있을 때 평가 대상에서 가장 큰 값을 가지고 있는 위치 값을 선택하기 때문이다. 이후 최고값을 갖는 여러 경로들을 대상으로 하여 깊이가 가장 큰 값을 갖는 경로를 추출한다. 이때 추출된 경로가 비교 기준 대상인 KEGG 신호 전달 경로와 일치하는 지에 대한 유사도 성능 평가를 실시한 다.

IV. 적용 및 평가

4.1 MAPK에 적용

MAP(Mitogen-Activated Protein) Kinase 신호전달경로는 세포막 단백질에 세포분열 유도물질인 Mitogen이 결합하며 시작된다. 이러한 신호전달 결과 세포 분화, 분열, 생존, 사망 등의 현상이 일어난다. 이러한 MAP Kinase 신호전달 경로는 세포막에서 발생한 신호를 핵 안쪽까지 증폭하면서 전달하고, 각 단계에서 어느 단백질을 활성화 하느냐에 따라 다양한 반응을 나타낼 수 있다. KEGG MAPK Signaling Pathway는 [그림 7]과 같다.

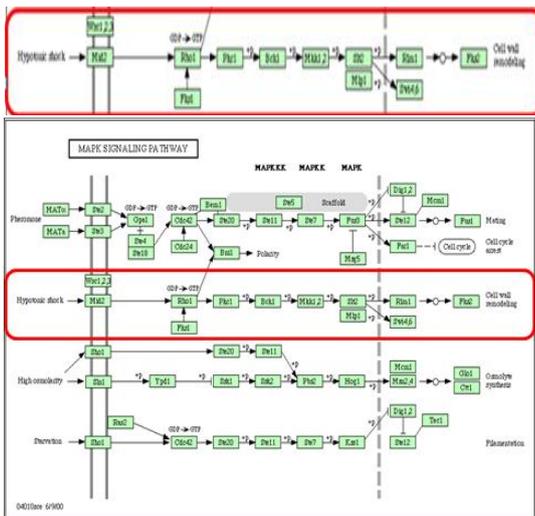


그림 7. MAPK 신호전달 경로
Fig 7. MAPK Signaling Pathway (S.cerevisiae)

KEGG MAPK Signaling Pathway는 4개의 Function으로 구분할 수 있다. 그 중 실험에서는 "Hypotonic shock" 기능에 대해 실험하여 그 결과를 분석한다.

4.2 평가 및 분석

알고리즘을 거쳐 산출된 평가 값을 내림차순 하여 가장 큰 값을 선정한 후 그 중 가장 많은 노드를 포함하고 있는 패스를 선정하여 비교 기준 대상인 KEGG에서 제공하는 MAPK "Hypotonic shock" 신호전달 경로를 실험 결과와 비교 하였다.

P	Q	R	S	T	U	V	W
start : YLR332W, target : YER111C, maximum length : 8							
YLR332W	YOL109W	YBR069C	YLR372W	YMR058W	YDL029W	YHR030C	YER111C
YLR332W	YOL109W	YBR069C	YLR372W	YER118C	YLR096W	YHR030C	YER111C
YLR332W	YOL109W	YBR069C	YLR372W	YER118C	YHL007C	YPL256C	YER111C
YLR332W	YOL109W	YBR069C	YLR372W	YDR456W	YGL008C	YHR030C	YER111C
YLR332W	YOL109W	YBR069C	YLR372W	YCR024C	YGL008C	YHR030C	YER111C
YLR332W	YOL109W	YBR069C	YLR372W	YGL054C	YNL201C	YPL153C	YER111C
YLR332W	YOL109W	YBR069C	YJR010C	YJR091C	YBR160W	YLR182W	YER111C

그림 8. 알고리즘 수행을 위한 입력 데이터
Fig 8. Input Data for LSPF Algorithm Practice

"Hypotonic shock" 기능 중 4개의 패스를 대상으로 알고리즘[15] 방법1과 방법2를 이용하여 신호전달 경로 추출 실험을 한 결과 [그림 9]와 같은 결과를 얻었다. 다시 말해 4개의 패스 중 2개의 패스는 목표노드 7개 중 4개 노드가 일치하는 패스를 찾았고 나머지 2개의 패스는 5개 노드가 일치하는 패스를 찾았다. 여기서의 특징은 LSPF 알고리즘 방법1과 방법2가 모두 일치하는 결과를 얻었다. 이 결과를 통해 알 수 있는 것은 방법1과 방법2 알고리즘 모두 단백질 위치정보 기반에서 착안한 알고리즘으로 방법이 매우 유사한 알고리즘이라 할 수 있겠다.

Function	Pathway							방법1.횟수	방법1.정확률	방법2.횟수	방법2.정확률
	1	2	3	4	5	6	7				
Hypotonic shock	YLR332W	YPR165W	YBL105C	YJL095W	YOR231W	YHR030C	YER111C	4	57%	4	57%
	YLR332W	YPR165W	YBL105C	YJL095W	YOR231W	YKR095W	YER111C	4	57%	4	57%
	YLR332W	YPR165W	YBL105C	YJL095W	YPL140C	YHR030C	YER111C	5	71%	5	71%
	YLR332W	YPR165W	YBL105C	YJL095W	YPL140C	YKR095W	YER111C	5	71%	5	71%

그림 9. "Hypotonic shock" 경로 및 수행결과
Fig 9. "Hypotonic shock" Pathway and Practice Result

[그림 9]의 결과를 살펴보면 실험 대상인 4개의 패스 중 2개의 패스는 비교 기준인 KEGG와의 유사도 측정에서 57%의 정확도를 얻었으며 나머지 2개의 패스는 71%의 정확도를 얻었다.

KEGG MAPK Signaling Pathway의 경로 즉 [그림 7]에서 제시된 경로들 중 시작 단백질과 목표 단백질이 같은 경로들은 유일한 경로 하나로 통일하여 프로그램 실행에 소요된 시간표를 [그림 10]에서 제시한다.

시작단백질	종료단백질	KEGG-Depth	본 System	Program_Start	Program_End	OutputFile_Size	소요시간
YLR332W	YPL089C	7	8	2008-11-04 4:23 AM	2008-11-04 4:38 PM	639k	12시간 15분
YLR332W	YER111C	7	8	2008-11-04 4:31 AM	2008-11-04 6:26 PM	1.15m	13시간 55분
YLR332W	YLR182W	7	8	2008-11-04 4:27 AM	2008-11-04 4:39 PM	2.13m	12시간 12분

그림 10. 신호전달 경로 패스 프로그램 실행 시간표
Fig 10. Signaling Pathway Program Practice Time

YLR332W를 시작 단백질로 하고 YER111C를 종료 단백질로 하는 실험에서 방법1 알고리즘을 적용한 결과 [그림 11]에 대해 분석하여 살펴 보자. 단백질 상호작용 데이터를

가지고 알고리즘을 적용하여 단백질 신호전달 경로를 찾는 결과와 목표로 하는 KEGG에서 찾고자 하는 7개의 패스 중 4개와 5개 각각의 패스를 정확하게 찾고 있으며, 방법2 알고리즘을 적용한 결과인 [그림 12]에 대해 분석하여 살펴 보면 방법1 알고리즘 적용의 경우와 동일한 결과를 얻었다.

단백질의 위치정보를 기반으로 한 실험은 초기 제안하는 연구로 생물학자나 컴퓨터 바이오인포매틱스를 연구하는 학자들에게 차기 단백질 신호전달 경로 연구에 기여될 것이라 사료된다.

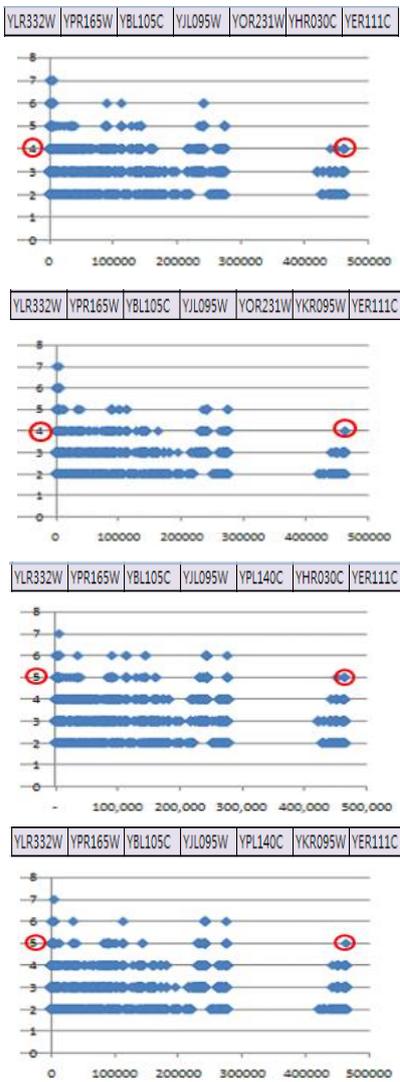


그림 11. 방법1 알고리즘 수행결과 (4 패스)
Fig 11. Practice Result of Method1 Algorithm (4 Path)

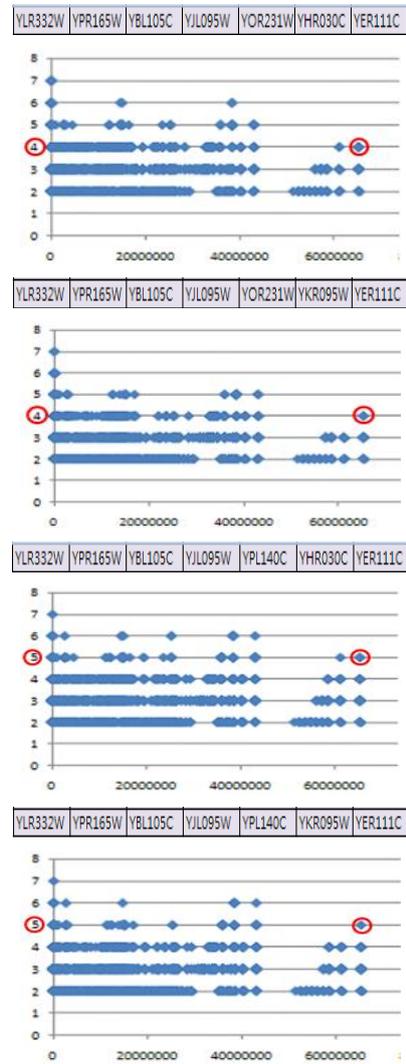


그림 12. 방법2 알고리즘 수행결과 (4 패스)
Fig 12. Practice Result of Method2 Algorithm (4 Path)

V. 결론 및 향후 계획

선행연구로는 신호전달 경로 추출에 있어서 두 하이브리드를 통해 얻은 단백질 상호작용 단백질과 DNA Microarray Data 실험으로부터 얻은 발현 프로파일을 사용 한다. 그러나 좋은 결과를 얻지 못했다. 본 연구에서는 Yeast 상호작용 데이터 46,121개의 단백질 정보와 약 5,000여 개의 Yeast (*S. cerevisiae*) 위치 정보를 이용하여 신호 전달 경로 추출을 하였다. 이때 상호작용을 기반으로 출발점을 세포막 단백질로

하고 도착점을 핵에 있는 단백질로 하였으며 신호전달 경로를 추출 하였다. 시뮬레이션 결과를 통하여, KEGG에서 제공하는 MAPK "Hypotonic shock" 기능의 신호전달 경로와 실험 결과에서 얻은 신호전달 경로의 유사도를 측정한 결과 방법1과 방법2 알고리즘(15) 모두 동일하게 4개의 패스 중 2개의 패스는 57%의 정확도를 얻었으며 나머지 2개의 패스는 71%의 정확도를 얻었다. 단백질의 경로를 찾아 내는 다양한 연구가 있었지만, 위치정보에 기반하여 경로를 찾으려는 시도는 "S. cerevisiae 단백질간 상호작용과 세포 내 위치 정보를 활용한 MAP Kinase 신호전달경로추출 및 예측을 위한 고성능 알고리즘 연구"(15)가 처음이다. 많은 생물학자나 컴퓨터공학자들 사이에 집중이 되는 연구이기는 하지만 아직까지 활발한 연구가 되지 못하는 이유 중에 하나는 생물학 관점에서 접근하였을 때 단백질 상호작용은 한곳에서 정지하며 단백질 간 상호작용이 일어나기도 하지만 때로는 이동하면서도 상호작용이 일어나기 때문에 신호전달 과정 중 잠시 존재했다가 사라지는 현상을 단백질 간 상호작용 결과로 밝혀내기 어려운 점으로 사료된다. 차기 연구에서는 적용범위 기능을 좀 더 확대하여 실험함으로써 실험 정확도를 더욱 높이고 아직 실험을 통해 밝혀 지지 않은 단백질들이 많음을 감안 할 때 미지의 단백질을 고려한 신호전달 경로 찾기와 여러 다양한 패스 적용에 대한 연구가 필요할 것으로 사료된다.

참고문헌

- [1] Schikowski B, Uetz P and Fields S, "A network of Protein-Protein Interaction in Yeast" Nat Biotechnol, 18:pp. 1257-1261, 2000
- [2] Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M and Pochart P, "A comprehensive analysis of Protein-Protein Interactions in Saccharomyces Cerevisiae" Nature, 2000
- [3] L. Giot, J. S. Bader, C. Brouwer et al., "A Protein Interaction Map of Drosophila melanogaster", Science, Vol. 302, No. 5651, pp. 1727-1736, 2003
- [4] Reuven Cohen, Shlomo Havlin, "Scale-Free Networks Are Ultrasmall" vol. 90, 90-94, 2003
- [5] Ibert-Laszlo arabasi, Zoltan N. Oltvai, "Understanding the Cell's Functional Organization", Nature, Vol. 5, pp. 101-103, 2004
- [6] Silvia D. M. Santos, Peter J. Verveer, Philippe I. H. Bastiaens, "Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate" Nature, Vol. 9, pp. 324-330, 2007
- [7] Jose B. Pereira-Leal, Anton J. Enright, Christos A. Ouzounis, "Detection of functional modules from protein interaction networks" Vol. 54, pp. 49-57, 2004
- [8] Victor Spirin, Leonid A. Mirny, "Protein complexes and functional modules in molecular networks" Vol. 100, No. 21, 12123-12128, 2003
- [9] 김은하, "단백질 상호작용 정보와 위치정보를 활용한 신호전달 경로 추출 및 시각화 시스템 구현" 이화여자대학교, 2006
- [10] DIP. <http://dip.doe-mbi.ucla.edu>
- [11] MIPS. <http://mips.gsf.de>
- [12] YPLD. <http://ypl.uni-graz.at/pages/home.html>
- [13] KEGG. <http://www.genome.jp/kegg/>
- [14] Won-ki Huh, James V. Falvo et al., "Global analysis of Protein localization in budding Yeast." Nature, 2003
- [15] 조미경, 김민경, 박현석, "S. cerevisiae 단백질간 상호작용과 세포 내 위치 정보를 활용한 MAP Kinase 신호전달경로추출 및 예측을 위한 고성능 알고리즘 연구", 한국컴퓨터정보학회 논문지, 제14권, 제3호, 193-207쪽, 2009년 3월.

저자 소개



조미경

소속 : 이화여자대학교
학위 : 이화여자대학교 공과대학 공학박사
관심분야 : Natural Language Processing
Bioinformatic
Genome Informatics



서정만

소속 : 한국재활복지대학교
컴퓨터게임개발과 부교수
학위 : 충북대학교 공학박사
관심분야 : 데이터베이스 시스템,
컴퓨터게임개발,
실시간처리



박현석

소속 : 이화여자대학교부교수
컴퓨터공학전공주임교수
컴퓨터전자공학부장
과학기술경영연계전공주임교수
학위 : Ph.D. (영)Univ. of Cambridge
(Computer Science and Engineering)
관심분야 : Natural Language Processing
Bioinformatics
Genome Informatics
Human Computer Interaction