

아토피관련 질병 네트워크로부터 질병단백체 발굴

Identification of Diseasomal Proteins from Atopy-Related Disease Network

이윤경*, 여명호**, 강태호**, 유재수**, 김학용*
충북대학교 생화학과*, 충북대학교 정보통신학과**

Yoon-Kyeong Lee(biokorea01@cbnu.ac.kr)*, Myeong-Ho Yeo(mhyeo@netdb.cbnu.ac.kr)**,
Tae-Ho Kang(thkang@netdb.cbnu.ac.kr)**, Jae-Soo Yoo(yjs@cbnu.ac.kr)**,
Hak-Yong Kim(hykim@cbnu.ac.kr)*

요약

본 연구는 질병과 관련이 있는 단백질들은 질병 네트워크를 형성함에 있어서 매우 중요한 인자로 작용할 가능성이 있다는 아이디어에서 출발한다. 우리는 Online Mendelian Inheritance in Man(OMIM)으로부터 아토피관련 43개 단백질 데이터베이스를 확보하고 이 단백질들과 상호작용하는 단백질 네트워크를 구축하였다. 아토피관련 단백질 네트워크를 바탕으로 질병 네트워크를 구축하였다. 질병 네트워크로부터 질병단백체인 CCR5, CCL11, 및 IL4R을 발굴하였는데, 이들 모두는 단백질 네트워크에서 허브 단백질로 작용하는 것들이다. 허브단백질은 세포에서 필수단백질로 작용하는 것으로 알려져 있는데, 본 연구에서는 허브단백질이면서 동시에 질병에서 매우 중요한 역할을 할 것으로 기대되는 질병단백체로 역할하고 있음을 확인하였다. 본 연구에서 소규모 아토피 관련 질병네트워크를 구축하여 분석하였지만, 여기에 제안한 질병네트워크 분석이 복잡한 인간 질병체계의 분자 기작 및 생물학적 진행과정을 이해하는데 실마리를 제공할 것으로 기대한다.

■ 중심어 : | 아토피 질병 | 단백질 네트워크 | 질병 네트워크 | 질병단백체 |

Abstract

In this study, we employed the idea that disease-related proteins tend to be work as an important factor for architecture of the disease network. We initially obtained 43 atopy-related proteins from the Online Mendelian Inheritance in Man (OMIM) and then constructed atopy-related protein interaction network. The protein network can be derived the map of the relationship between different disease proteins, denoted disease interaction network. We demonstrate that the associations between diseases are directly correlated to their underlying protein-protein interaction networks. From constructed the disease-protein bipartite network, we derived three diseasomal proteins, CCR5, CCL11, and IL4R. Although we use the relatively small subnetwork, an atopy-related disease network, it is sufficient that the discovery of protein interaction networks assigned by diseases will provide insight into the underlying molecular mechanisms and biological processes in complex human disease system.

■ keyword : | Atopy | Protein Network | Disease Network | Diseasomal Protein |

* 본 논문은 2008년 교육과학기술부와 한국산업기술재단의 "지역혁신인력양성사업"과 교육과학기술부와 학술진흥재단의 "지역거점연구단육성사업/충북BIT연구중심대학육성사업단"의 지원을 받아 수행된 연구임.

접수번호 : #081210-003

심사완료일 : 2009년 03월 05일

접수일자 : 2008년 12월 10일

교신저자 : 김학용, e-mail : hykim@cbnu.ac.kr

I. 서론

인간 질병에 관여하는 단백질들은 질병의 병리학적 기작 및 임상학적 실재를 이해하는데 중요한 요소이다. 전통적인 방법인 한 질병에 관여하는 단백질과의 연결고리를 분석하는 제한된 연구를 통해서 질병사이의 복잡한 상관관계를 분석하고 이해하는 것은 한계가 있다. 지난 몇 년 동안 척도없는(scale free) 네트워크 모델을 통해서 생체 시스템에서 발견되는 많은 대사체, 유전체, 및 단백질체과 같은 세포 구성성분들 간의 상관관계를 이해하였다[1]. 척도없는 네트워크는 링크수가 많은 소수의 허브 노드와 링크수가 적은 다수의 노드들이 공존하는 구조를 취하고 있다. 네트워크에서 일반적으로 다른 단백질에 비해 상호작용이 많은 단백질을 허브(hub) 단백질이라 한다[2]. 허브 단백질은 네트워크에서 경로를 단축시키는 역할을 하지만 동시에 세포 생존에 중요하게 관여하는 필수(essential) 단백질 혹은 치사(lethal) 단백질일 확률이 높다고 보고되었다[3][4].

네트워크의 구조적 및 기능적 분석방법과 질병관련 네트워크와의 결합을 통해서 복잡한 인간 질병체계를 이해하고자 한다면 시스템수준에서 질병을 이해하는 새로운 통찰력을 도출할 수 있을 것이다[5]. 지난 수십 년 동안, 의학적 임상학적 접근방법은 질병과 한 두 개의 단백질사이의 상관관계를 분석함으로써 질병을 이해하고자 하였다. 최근에 하나의 질병을 분석하기 보다는 질병사이의 상관관계를 보여주는 질병 네트워크의 구조와 양상을 분석하고자하는 새로운 접근방법이 제안되었다. 최근에 인간 질병과 유전자 사이에 상관관계를 규명하는 연구 결과가 보고되었는데, 하나의 유전 질환은 다양한 유전자와 연결되어 있고 이들 유전자는 또 다른 질환과 연결되어 결국 비슷한 질병들이 연결고리를 형성하고 있음을 밝혔다[6].

비록 유전 질환일지라도 표현형으로 나타나는 것은 단백질이다. 다시 말해 유전자가 병인의 근원은 되지만 실제 질환의 발생, 작용 기작 등은 표현형인 단백질에 의해서 좌우된다. 병의 원인을 규명하는 것은 유전자로부터 접근하는 것이 바람직하지만 질병의 확산, 질병과의 상관관계 등을 이해하는 데는 단백질을 분석하는 것

이 바람직하다. 특히, 단백질 분자는 각각 서로 다른 분자를 인지할 수 있고, 적절한 상황에서 서로 결합하여 공동으로 기능을 수행한다[7]. 따라서 복잡한 질병과의 상관관계의 분석은 단백질 네트워크를 분석하는 것이 필수적이라 하겠다. 본 연구에서는 아토피 관련 질병 단백질 상호작용 네트워크로부터 질병 상호작용 네트워크를 도출하고 이 네트워크의 분석을 통해 아토피 관련 질병단백체를 발굴하고자 한다. 나아가서 이 질병단백체의 3차 구조, 활성부위 등 규명하면 아토피 관련 질환의 저해 또는 통제 할 수 있는 선도화합물의 단초를 제공할 수 있을 것으로 기대한다.

II. 관련 연구

세포를 조절하는 생화학 경로가 무너지면 관련 질병으로 이어질 수 있다. 같은 맥락에서 단백질 상호작용 네트워크가 무너진다면 질병 사이의 연결고리가 무너지면서 여러 질병이 생길 수 있다[8]. 왜냐하면 단백질 상호작용 네트워크는 필연적으로 분자네트워크를 형성하며 생명체 기능의 항상성을 유지하고 있기 때문이다.

단백질 네트워크로부터 기능이 밝혀지지 않은 단백질을 예측할 수 있듯이[7], 네트워크상에서 단백질은 다른 단백질과 매우 밀접하게 연결되어 있고 유사한 질병에 관여하기 때문에 밝혀지지 않은 질병관련 단백질을 예측할 수 있을 것이다[9]. 예를 들어, 헌팅톤병(Huntington's disease) 관련 단백질 네트워크에는 헌팅톤 관련 다수의 단백질이 상호작용하고 있는데, 그 중에서 G protein coupled receptor kinase interaction protein인 GIT1가 이 질병에 관여하는 새로운 단백질이 확인되었다[10].

최근 미생물 병원체의 단백질 상호작용 네트워크에 관한 연구가 증가하면서 네트워크를 통한 전염병 연구 분야가 주목을 받고 있다[11]. 인간 단백질 네트워크와 병원체 네트워크를 비교함으로써 공통 상호작용 단백질을 추출할 수 있으며 이 단백질은 신약개발을 위한 자료로 활용할 수 있다. 병원체뿐만 아니라 바이러스 단백질 네트워크와 사람 단백질 네트워크와의 비교를

통해 전염성 질환에 대한 병인, 기작, 치료 등의 연구가 확산되고 있다[12]. 이는 궁극적으로는 인간과 환경과의 상관관계까지도 분석할 수 있기 때문에 이러한 단백질 네트워크에 대한 연구는 급속히 팽창할 것이다.

III. 연구 방법 및 연구 결과

1. 연구방법

아토피 관련 43개 단백질을 OMIM (Online Mendelian Inheritance in Man) 데이터베이스로부터 2008년 4월에 얻었다[13]. 본 연구를 마친 다음 조사한 결과 AAD, RAG, KTP, KTPD, HLA-DRB1 등 5개 단백질이 새로이 추가되어 48개 아토피 관련 단백질이 존재하나 본 연구에서는 이들 단백질을 제거한 43개 단백질을 가지고 연구를 수행하였다. 이 아토피 관련 단백질은 [표 1]에 요약하였다.

단백질-단백질 상호작용 데이터베이스는 HPRD (Human Protein Reference Database)로부터 얻어[14] 아토피관련 단백질 네트워크를 구축하고 InterViewer 프로그램 (인하대 바이오컴퓨팅연구실)으로 시각화하였다. 네트워크에 포함된 단백질들의 질병정보를 OMIM으로부터 확보하여 질병 명으로 치환함으로써 질병단백질을 구축하였다.

네트워크는 노드 (단백질)와 링크 (단백질 상호작용)를 결합시킨 것으로 기본 원리는 다음과 같다. 만약 어떤 단백질이 다른 단백질과 결합하는 경우 값 1을 부여하고, 어떤 단백질이 다른 단백질과 전혀 결합하지 않는 경우 값을 0을 부여하여 구축하였으며 이때 방향성은 고려하지 않았다[7].

표 1. 43개 아토피 관련 단백질 데이터

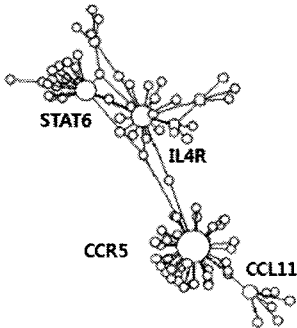
번호	단백질명	Gene Map Locus
1	IGE RESPONSIVENESS	16p11
2	CALCIUM-BINDING ATOPY-RELATED AG 1	-
3	INTERLEUKIN 4 RECEPTOR	16p12.1-p11.2
4	MEMBRANE-SPAN, 4 DOMAINS, SUBFAMILY A, MEMB, 2	11q13
5	HEPATITIS A VIRUS CELLULAR RECEPTOR 1	5q33.2

6	PHOSPHOLIPASE A2, GROUP VII	6p21.2-p12
7	SELECTIN P	1q23-q25
8	SERINE PROTEASE INHIBITOR, KAZAL-TY, 5	Chr.X
9	IMMUNOGLOBULIN A DEFICIENCY 1	6p21.3
10	SQUAM. CELL CARCINOMA Ag BY T CELLS 1	Chr.11
11	IMMUNODYSREGULATION, POLYENDOCRINOPATHY	Xp11.23-q13.3
12	INTERLEUKIN 13	5q31
13	PHD FINGER PROTEIN 11	13q14.1
14	DERMATITIS, ATOPIC, 6	5q31-q33
15	CHEMOKINE, CC MOTIF, RECEPTOR 5	3p21
16	ASTHMA-RELATED TRAITS, SUSCEPTIBILITY TO, 2	1q32.1
17	ASTHMA-RELATED TRAITS, SUSCEPTIBILITY TO, 6	17q21
18	ASTHMA-RELATED TRAITS, SUSCEPTIBILITY TO, 4	1p31
19	ASTHMA-RELATED TRAITS, SUSCEPTIBILITY TO, 3	2p16
20	ASTHMA-RELATED TRAITS, SUSCEPTIBILITY TO, 1	14q22.1
21	DERMATITIS, ATOPIC	13q12-q14
22	CHEMOKINE, CC MOTIF, LIGAND 24	7q11.23
23	SIGNAL TRANSD. AND ACTIVATOR OF TRANS, 6	12q13
24	CHEMOKINE, CC MOTIF, LIGAND 11	17q21
25	TUMOR PROTEIN	13q12-q14
26	UTEROGLOBIN	11q12.3-q13.1
27	INTERLEUKIN 12B	5q31.1-q33.1
28	KERATOCONUS 1	20p11.2
29	IMMUNOGLOBULIN Gm2	14q32.33
30	FILAGGRIN	1q21
31	CHYMASE 1	14q11.2
32	HUMAN IMMUNODEFICIENCY VIRUS TYPE 1	19q13.4
33	CASPASE RECRUITMENT DOMAIN-CONTAINING PROTEIN 11	-
34	INTERLEUKIN 13 RECEPTOR, ALPHA-1	Chr.X
35	INTERLEUKIN 9	5q31.1
36	HISTOCOMPATIBILITY COMPLEX, CLASS II, DR BETA-1	6p21.3
37	INTERLEUKIN 10	1q31-q32
38	GAP JUNCTION PROTEIN, BETA-2	13q11-q12
39	PROTEASE INHIBITOR 1	14q32.1
40	ALLERGIC RHINITIS	5q31
41	ASTHMA-RELATED TRAITS, SUSCEPTIBILITY TO, 5	12q14.3
42	ICHTHYOSIS PREMATURITY SYNDROME	-
43	ASTHMA-RELATED TRAITS, SUSCEPTIBILITY	17q21

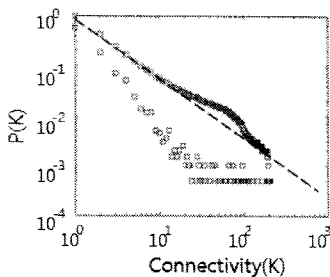
2. 연구결과

[표 1]에서 보여준 질병관련 단백질과 HPRD에서 얻은 상호작용 데이터를 활용하여 아토피 관련 단백질 상호작용 네트워크를 구축하였다. 구축된 아토피 관련 단백질 상호작용 네트워크는 147개 노드(단백질)와 144개 링크(상호작용)로 구성되었으며 이 들 중에서 가장 큰 네트워크는 87개의 단백질과 94개의 상호작용으로 만들어졌다[그림 1A]. 비록 아토피 관련 소규모 네트워크를 구축하였지만 전형적인 측도없는 네트워크[그림 1B] 및 계층적 네트워크특성을 보여주었으며[그림 1C], STAT6, IL4R, CCR5, CCL11 등과 같은 허브 단백질도 발견되었다. 네트워크를 구성하는 단백질을 질병으로 전환하였는데, 만약 단백질 중에서 다수의 질병과 관련이 있으면 하나의 단백질이 다수의 질병 노드로 전환되며 단백질 중에서 질병과 관련이 없는 단백질은 전환과정에서 제거된다. 이 과정을 통해서 아토피관련 질병 네트워크를 구축하였다[그림 2].

A. 단백질 네트워크



B. Connectivity에 따른 단백질 수 변화



C. Connectivity에 따른 뭉침계수 변화

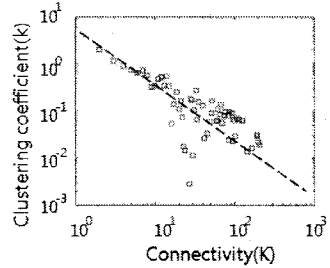


그림 1. 아토피 관련 단백질-상호작용 네트워크

24개 질병 중에서 천식, 아토피, HIV, 및 CCR polymorphism등이 비교적 많은 링크를 가지고 있었다. 예상한바와 같이, 면역관련 단백질이 관여하는 아토피와 천식 및 chemokine 수용체를 공유하는 HIV와 CCR polymorphism이 밀접한 관련성이 있음을 보여주고 있다. 아토피 및 천식이 HIV와 밀접한 관계가 있음을 보여주었는데, 이들 질병이 면역과 관련되어 있기 때문에 해석하면 충분히 예견되기도 하지만, 현 시점에는 명확한 임상학적 정보가 없으나 만약 아토피나 천식 환자가 HIV에 더욱 민감하다는 임상학적인 결과가 나타난다면 우리는 구축한 네트워크로부터 매우 중요한 결과를 도출하였다고 볼 수 있다.

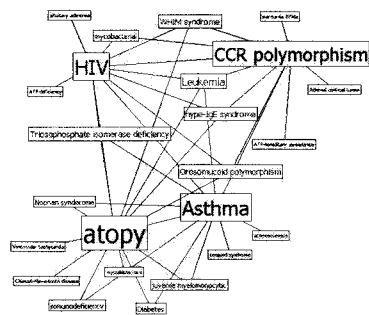


그림 2. 아토피 관련 질병 네트워크

질병단백체를 발굴하기 위하여 질병과 질병관련 단백질 사이의 상관관계를 보여주는 세트를 구축하였다 [그림 3]. 각 질병에 관여하는 단백질과 질병과의 연결

고리를 통해서 우리는 각 질병들에게 많이 발견되는 질병단백체를 발굴할 수 있다[6]. [그림 1]에서 보여준 4개의 허브 단백질 중에서 CCR5, CCL11, IL4R이 질병과 매우 밀접한 질병단백체임을 확인하였다. 이는 허브 단백질이 일반적으로 치사 단백질 또는 필수 단백질일 가능성이 높듯이 질병네트워크에서도 매우 중요한 역할을 하고 있음을 의미한다.

STAT6가 허브 단백질임에도 불구하고[그림 1], 질병단백체는 아닌 것으로 나타났다. STAT6는 cytokine이나 성장인자에 의해서 활성화되는 수용체와 연계된 카나제에 의해 인산화되는 단백질로 신호전이 경로에 매우 중요한 역할을 한다[15]. 비록 STAT6이 아토피 질병관련 신호경로를 조절하지만, 여러 질병이 연결고리를 형성하는 비교적 복잡한 질병계인 질병 네트워크에는 관여하지 않는 것으로 보인다. STAT6와 상호작용하는 단백질들은 면역관련 질병과 상관관계가 부족한 것으로 추정된다.

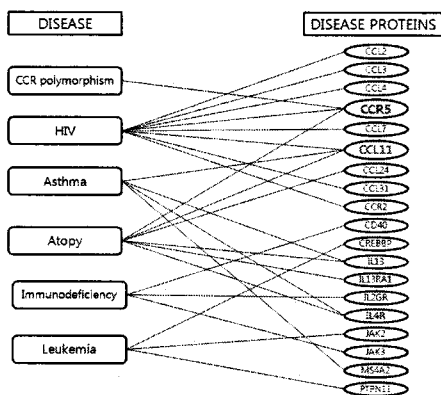


그림 3. 아토피 질병 네트워크로부터 질병단백체 발굴

Chemokine 수용체인 CCR5는 T 세포에 많이 존재하기 때문에 천식과 아토피와 같은 면역계 질환과 관련이 있을 것이다. HIV는 CCR5를 이용하여 목표 세포에 침투하는 것으로 알려졌다[16]. 따라서 CCR5가 HIV가 포함된 질병 네트워크에서 질병단백체인 것은 당연한 결과일 것이다. Cytokine인 CCL11은 CCR5의 리간드이기 때문에 HIV와 연관이 있을 수 있다[17]. 그러나 아직까지 CCL11이 HIV와 연계되어 있다는 어떤 증거도 보고된

바가 없다. 생물정보학적 접근방법으로 얻은 이 결과가 앞으로 있을 실제 생물학적 실험결과와 일치된다면 매우 의미가 있을 것이다.

IV. 결론

본 연구에서 우리는 생물정보학적 접근 방법을 이용하여 질병 관련 단백질 네트워크와 질병 네트워크와의 상관관계를 규명하였다. 간략하게 요약하면, OMIM 데이터베이스로부터 아토피 관련 43개 단백질을 얻은 다음[표 1], 단백질 상호작용 네트워크를 구축하였다[그림 1]. 네트워크에 포함된 단백질을 관련 질병으로 전환하여 질병 단백질 네트워크를 구축한 다음 [그림 2], 이를 바탕으로 질병단백체를 발굴하였다[그림 3]. 이러한 접근 방법은 질병-질병 상관관계를 분석뿐만 아니라 질병사이의 분자적 상관관계를 보여줄 수 있다[8]. 우리는 이를 토대로 질병단백체인 CCR5, IL4R, 및 CCL11을 발굴하였으며 이들 단백질이 질병 네트워크에서 매우 밀접하게 연결되고 있음을 확인하였다.

이러한 접근 방법을 통해 발굴한 질병단백체는 분자 수준에서 선도화합물에 대한 실마리를 제공할 수 있을 것이다. 네트워크를 통해 밝혀진 잘 알려진 질병과 알려지지 않은 질병사이의 상관관계를 통해 분자 수준에서 연결고리를 확인할 수 있다. 본 연구에서 아토피나 천식이 HIV와 관련이 있다는 결과는 하나의 질병을 대상으로 하는 단순계 수준에서 접근하였다면 불가능하였을 것이다.

결론적으로, 비록 전체 질병 시스템이 아닌 소규모 아토피 관련 질병 네트워크를 분석하였지만, 본 연구에서 제시한 생물정보학적 접근 방법은 좀 더 복잡한 질병시스템을 분자 수준에서의 기작을 이해하는데 깊은 통찰력을 제공할 것이다[18]. 또한 사람 질병 단백질 전체 네트워크를 구축하여 질병단백체를 확보하고 이미 노산 서열이나 도메인의 특성을 분석하여 신약개발을 위한 선도화합물 발굴하는데 정보를 제공할 수 있을 것으로 기대한다.

참고 문헌

- [1] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. -L. barabasi, "The large-scale organization of metabolic networks," *Nature*, Vol.407, No.6804, pp.41-42, 2001.
- [2] A. -L. Barabasi and Z. N. Oltvai, "Network biology: understanding the cell's functional organization," *Nat. Rev. Genet.*, Vol.5, No.2, pp.101-113, 2004.
- [3] H. Jeong, S. P. Mason, A. L. Barabasi, and Z. N. Oltvai. "Lethality and centrality in protein networks," *Nature*, Vol.411, No.6833, pp.41-42, 2001.
- [4] H. Yu, D. Greenbaum, H. X. Lu, X. Zhu, and M. Gerstein. "Genomic analysis of essentiality within protein networks," *TRENDS in Genomics*, Vol.20, No.6, pp.227-231, 2004.
- [5] S. Bortoluzzi, C. Romualdi, A. Bisognin, and G. A. Danieli, "Disease genes and intracellular protein networks," *Physiol. Genom.* Vol.15, No.3, pp.223-227, 2003.
- [6] K.-I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A. -L. Barabasi, "The human disease network," *Proc. Nat'l. Acad. Sci.* Vol.104, No.21, pp.8685-8690, 2007.
- [7] J. W. Ryu, H. Y. Kim, T. -H. Kang, and J. -S. Yoo, "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] L. Sam, Y. Liu, J. Li, C. Friedman, and Y. A. Lussier, "Discovery of protein interaction networks shared by diseases," *Pacific Sym. on Biocomp.* Vol.12, No.1 pp.76-87, 2007.
- [9] T. Ideker and R. Sharan, "Protein networks in disease," *Genome Res.* Vol.18, No.4, pp.644-652, 2008.
- [10] H. Goehler, M. Lalowski, U. Stelzl, S. Waelter, M. Stroedick, U. Worm, and E. E. Wanker, "A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease," *Mol. Cell.* Vol.15, No.6, pp.853-865, 2004.
- [11] J. R. Parrish, J. Yu, G. Liu, J. A. Hines, J. E. Chan, and R. L. Finlev Jr., "A proteome-wide protein interaction map for *Campylobacter jejuni*," *Genome Biol.* Vol.8, No.7, p.R130, 2007.
- [12] M. A. Calderwood, K. Venkatesan, L. Xing, M. R. Chase, A. Vazquez, A. M. Holthaus, E. Kieff, and E. Johannsen, "Epstein-Barr virus and virus human protein interaction maps," *Proc. Nat'l. Acad. Sci.*, Vol.104, No.18, pp.7606-7611, 2007.
- [13] A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick, "Online mendelian inheritance in man(OMIM), a knowledgebase of human genes and genetic disorders," *Nucleic Acids Res.* Vol.33, No.DB, pp.D514-D517, 2005.
- [14] T. S. K. Prasad, R. Goel, K. Kandasamy, S. Keerthikumar, S. Kumar, P. Chaerkady, and A. Pandey, "Human protein reference database-2009 update," *Nucleic Acids Res.*, Vol.36, No.dbi, pp.1-6, 2008.
- [15] C. R. Yu, R. M. Mahdi, S. Ebong, B. P. Vistica, J. Chen, Y. Guo, I. Gery, and C. E. Eqwuaqu, "Cell proliferation and STAT6 pathways are negatively regulated in T cells by STAT1 and suppressors of cytokine signaling," *J. Immunol.* Vol.173, No.2, pp.737-746, 2004.
- [16] M. M. Lederman, A. Penn-Nicholson, M. Cho, and D. Mosier, "Biology of CCR5 and its role in HIV infection and treatment," *J. of Am. Med. Asso.*, Vol.296, No.7, pp.815-826, 2006.
- [17] P. Ogvie, G. Bardi, I. Clark-Lewis, M.

