



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2015년01월30일
(11) 등록번호 10-1488338
(24) 등록일자 2015년01월26일

| | |
|---|--|
| <p>(51) 국제특허분류(Int. Cl.) G06F 19/10 (2011.01)</p> <p>(21) 출원번호 10-2014-0141721</p> <p>(22) 출원일자 2014년10월20일 심사청구일자 2014년10월20일</p> <p>(56) 선행기술조사문헌 KR101243063 B1 바이오패스웨이를 위한 지식 표현 시스템(정보과학회논문지:소프트웨어및응용제31권제3호,(2004.03)) KR1020140016784 A JP2003167882 A</p> | <p>(73) 특허권자 한국과학기술정보연구원 대전광역시 유성구 대학로 245 (어은동)</p> <p>(72) 발명자 최윤수 대전광역시 서구 청사로 65, 103동 1303호 (월평동, 황실타운아파트)</p> <p>전선희 대전광역시 중구 대흥로 176-15, 6동 405호 (대흥동, 현대아파트) (뒷면에 계속)</p> <p>(74) 대리인 김용인, 지관영</p> |
|---|--|

전체 청구항 수 : 총 20 항

심사관 : 조현경

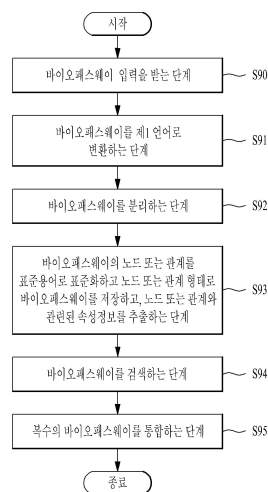
(54) 발명의 명칭 **바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체**

(57) 요약

본 발명은 다양한 바이오패스웨이 관련 기관에서 다양한 바이오패스웨이 기술 언어(KGML, SBML, BIOPAX 등)를 사용하여 바이오패스웨이를 개별적 독립적으로 구축하기 때문에 이질적인 언어로 표현된 바이오패스웨이를 효과적으로 통합하기 위함이다.

이를 위해 본 발명은, 바이오패스웨이 입력을 받는 단계(S90), 바이오 패스웨이를 제1언어로 변환하는 단계(S91), 바이오패스웨이를 분리하는 단계(S92), 바이오패스웨이의 노드 또는 관계를 표준용어로 표준화하고 노드 또는 관계 형태로 바이오패스웨이를 저장하고, 노드 또는 관계와 관련된 속성정보를 추출하는 단계(S93), 바이오 패스웨이를 검색하는 단계(S94), 복수의 바이오패스웨이들을 통합하는 단계(S95)를 포함한다.

대표도 - 도9



(72) 발명자

서동민

대전광역시 서구 청사서로 29, 102동 901호 (월평동, 셋별아파트)

박동인

서울특별시 서초구 나루터로 46, 101동 503호 (잠원동, 금호베스트빌아파트)

유석중

충청북도 청주시 흥덕구 2순환로1375번길 20, 508동 602호 (가경동, 대우푸르지오아파트)

이민호

대전광역시 서구 관저동로90번길 47, 1109동 504호 (관저동, 느리울마을아파트11단지)

성원경

대전광역시 유성구 엑스포로 448, 102동 604호 (전민동, 엑스포아파트)

특허청구의 범위

청구항 1

적어도 하나 이상의 바이오패스웨이 입력을 받는 입력모듈;
 상기 적어도 하나 이상의 바이오패스웨이를 제1언어로 변환하는 언어통일모듈;
 상기 제1언어로 변환된 바이오패스웨이를 분리하는 분리모듈,
 상기 분리모듈은 상기 제1언어로 변환된 바이오패스웨이를 노드 또는 관계로 분리하는 것을 포함하고;
 상기 노드 또는 관계를 표준용어로 표준화하고 상기 노드 또는 관계의 형태로 상기 제1언어로 변환된 바이오패스웨이를 저장하고, 상기 노드 또는 관계와 관련된 속성정보를 추출하는 표준화모듈; 및
 복수의 상기 저장된 바이오패스웨이들을 통합하는 통합모듈을 포함하는 바이오패스웨이 통합장치.

청구항 2

제 1항에 있어서, 상기 바이오패스웨이 통합장치는,
 노드, 관계, 바이오패스웨이 일부분, 바이오패스웨이 전체를 이용하여 상기 저장된 바이오패스웨이를 검색하는 검색모듈을 더 포함하는 바이오패스웨이 통합장치.

청구항 3

제 1항에 있어서, 상기 적어도 하나 이상의 바이오패스웨이는,
 이질적인 언어로 구축된 네트워크 형태로 표현되는 것을 포함하는 바이오패스웨이 통합장치.

청구항 4

제 1항에 있어서, 상기 언어통일모듈은,
 상기 변환된 바이오패스웨이를 저장하는 것을 포함하는 바이오패스웨이 통합장치.

청구항 5

제 1항에 있어서, 상기 분리된 바이오패스웨이는,
 노드-관계-노드 형태로 분리하는 것을 포함하는 바이오패스웨이 통합장치.

청구항 6

제 1항에 있어서, 상기 표준화모듈은,
 노드-관계-노드와 같이 트리플 형태로 저장하는 것을 포함하는 바이오패스웨이 통합장치.

청구항 7

제 1항에 있어서, 상기 추출된 속성정보는,
 상기 입력된 적어도 하나 이상의 바이오패스웨이의 노드별 빈도수, 노드별 링크수, 관계별 빈도수, 상기 저장된 바이오패스웨이의 전체 노드별 빈도수, 전체 노드별 링크수, 전체 관계별 빈도수 및 이에 기반한 가중치 정보를 포함하는 바이오패스웨이 통합장치.

청구항 8

제 1항에 있어서, 상기 통합모듈은,
 사용자 지정에 의해 또는 검색에 의해 선택된 복수의 바이오패스웨이를 통합하는 것을 포함하는 바이오패스웨이 통합장치.

청구항 9

제 1항에 있어서, 상기 통합모듈은,
상기 통합된 바이오패스웨이를 저장하는 것을 포함하는 바이오패스웨이 통합장치.

청구항 10

제 1항에 있어서, 상기 통합모듈은,
바이오패스웨이가 관계정보를 기반으로 통합되는 것을 포함하는 바이오패스웨이 통합장치.

청구항 11

제 1항에 있어서, 상기 바이오패스웨이 통합장치는,
바이오패스웨이 DB와 정보 DB를 더 포함하는 바이오패스웨이 통합장치.

청구항 12

적어도 하나 이상의 바이오패스웨이 입력을 받는 단계;
상기 적어도 하나 이상의 바이오패스웨이를 제1언어로 변환하는 단계;
상기 제1언어로 변환된 바이오패스웨이를 분리하는 단계,
상기 분리하는 단계는 상기 제1언어로 변환된 바이오패스웨이를 노드 또는 관계로 분리하는 것을 포함하고;
상기 노드 또는 관계를 표준용어로 표준화하고 상기 노드 또는 관계의 형태로 상기 제1언어로 변환된 바이오패스웨이를 저장하고, 상기 노드 또는 관계와 관련된 속성정보를 추출하는 단계; 및
복수의 상기 저장된 바이오패스웨이들을 통합하는 단계를 포함하는 바이오패스웨이 통합방법.

청구항 13

제 12항에 있어서, 상기 바이오패스웨이 통합방법은,
노드, 관계, 바이오패스웨이 일부분, 바이오패스웨이 전체를 이용하여 상기 저장된 바이오패스웨이를 검색하는 검색하는 단계를 더 포함하는 바이오패스웨이 통합방법.

청구항 14

제 12항에 있어서, 상기 적어도 하나 이상의 바이오패스웨이를 제1언어로 변환하는 단계는,
상기 변환된 바이오패스웨이를 저장하는 단계를 더 포함하는 바이오패스웨이 통합방법.

청구항 15

제 12항에 있어서, 상기 제1언어로 변환된 바이오패스웨이를 분리하는 단계는,
상기 제1언어로 변환된 바이오패스웨이를 노드-관계-노드 형태로 분리하는 것을 포함하는 바이오패스웨이 통합방법.

청구항 16

제 12항에 있어서, 상기 노드 또는 관계를 표준용어로 표준화하고 상기 노드 또는 관계의 형태로 상기 제1언어로 변환된 바이오패스웨이를 저장하고, 상기 노드 또는 관계와 관련된 속성정보를 추출하는 단계는,
상기 표준화된 바이오패스웨이를 노드-관계-노드와 같이 트리플형태로 저장하는 단계를 더 포함하는 바이오패스웨이 통합방법.

청구항 17

제 12항에 있어서, 상기 추출된 속성정보는,

상기 입력된 적어도 하나 이상의 바이오패스웨이의 노드별 빈도수, 노드별 링크수, 관계별 빈도수, 상기 저장된 바이오패스웨이의 전체 노드별 빈도수, 전체 노드별 링크수, 전체 관계별 빈도수 및 이에 기반한 가중치 정보를 포함하는 바이오패스웨이 통합방법.

청구항 18

제 12항에 있어서, 상기 복수의 상기 저장된 바이오패스웨이를 통합하는 단계는, 사용자 지정에 의해 또는 검색에 의해 선택된 복수의 바이오패스웨이를 통합하는 것을 포함하는 바이오패스웨이 통합방법.

청구항 19

제 12항에 있어서, 상기 복수의 상기 저장된 바이오패스웨이를 통합하는 단계는, 바이오패스웨이가 관계정보를 기반으로 통합되는 것을 포함하는 바이오패스웨이 통합방법.

청구항 20

적어도 하나 이상의 바이오패스웨이 입력을 받고, 상기 적어도 하나 이상의 바이오패스웨이를 제1언어로 변환하고, 상기 제1언어로 변환된 바이오패스웨이를 노드 또는 관계로 분리하고, 상기 노드 또는 관계를 표준용어로 표준화하고 상기 노드 또는 관계의 형태로 상기 제1언어로 변환된 바이오패스웨이를 저장하고, 상기 노드 또는 관계와 관련된 속성정보를 추출하고, 복수의 상기 저장된 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체.

명세서

기술분야

[0001] 본 발명은 바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체에 관한 것이다.

배경기술

[0002] 바이오 인포매틱스는 생명공학(Bio)과 정보공학(Informatics)의 합성어로 컴퓨터와 소프트웨어를 활용해 유전자의 염기서열 데이터를 수집, 분석, 관리 및 활용하는 것으로 생물정보학이라고도 한다. 최근 인간 게놈 프로젝트 완적으로 유전 정보량이 증가하면서 관리할 데이터 양도 늘어나 생물정보학의 중요성이 높아졌다. 이와 같이 생물정보학의 발전과 함께 생물 관련 정보들이 기하급수적으로 증가하고 있으며 연구 대상도 DNA, RNA, 단백질에서 더 나아가 이들의 상호작용 및 조절 메커니즘에 의해 기능들이 어떻게 수행되는 지에 관한 바이오패스웨이 까지 포함하게 되었다. 바이오패스웨이는 광대한 양의 정보를 포괄하며 구성체 사이의 유기적 관계를 나타내고 있는 것이다. 생물학 분야에서 양질의 바이오패스웨이 데이터베이스는 다양한 생물체의 생명 활동 메커니즘 이해, 질병의 발병, 진행, 자연소멸 및 치유에 관한 실체적 원인규명 등에 활용될 수 있다. 생명공학 분야에서의 효율적인 연구개발과 더불어 지식서비스 관점에서의 실질적인 장점이 있지만, 다양한 바이오패스웨이 관련 기관에서 다양한 바이오패스웨이 기술 언어(KGML, SBML, BIOPAX 등)를 사용하여 개별적 독립적으로 구축하기 때문에 현재 바이오패스웨이 데이터베이스 구축, 연계, 활용 측면에서 많은 문제점과 한계점이 존재한다. 또한 현재 대부분의 바이오패스웨이 데이터베이스는 수작업으로 구축되므로 막대한 구축 비용이 필요하고 기술 발전에 맞춘 신속한 데이터베이스 확장 및 변경이 불가능하다.

발명의 내용

해결하려는 과제

[0003] 본 발명은 상기한 바와 같은 문제점을 해결하기 위한 것으로서 이질적인 언어로 표현된 바이오패스웨이를 효과적으로 통합하기 위한 바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체를 제공하는 데 그 목적이 있다.

과제의 해결 수단

[0004] 이와 같은 목적을 달성하기 위한, 본 발명의 제 1 측면에 따르면, 본 발명에 따른 바이오패스웨이 통합 장치는, 적어도 하나 이상의 바이오패스웨이 입력을 받는 입력모듈; 상기 적어도 하나 이상의 바이오패스웨이를 제1언어로 변환하는 언어통일모듈; 상기 제1언어로 변환된 바이오패스웨이를 분리하는 분리모듈, 상기 분리모듈은 상기 제1언어로 변환된 바이오패스웨이를 노드 또는 관계로 분리하는 것을 포함하고; 상기 노드 또는 관계를 표준용어로 표준화하고 상기 노드 또는 관계의 형태로 상기 제1언어로 변환된 바이오패스웨이를 저장하고, 상기 노드 또는 관계와 관련된 속성정보를 추출하는 표준화모듈; 및 복수의 상기 저장된 바이오패스웨이를 통합하는 통합모듈을 포함하는 패스웨이 통합 장치를 포함한다.

[0005] 본 발명의 제 2 측면에 따르면, 적어도 하나 이상의 바이오패스웨이 입력을 받는 단계; 상기 적어도 하나 이상의 바이오패스웨이를 제1언어로 변환하는 단계; 상기 제1언어로 변환된 바이오패스웨이를 분리하는 단계, 상기 분리하는 단계는 상기 제1언어로 변환된 바이오패스웨이를 노드 또는 관계로 분리하는 것을 포함하고; 상기 노드 또는 관계를 표준용어로 표준화하고 상기 노드 또는 관계의 형태로 상기 제1언어로 변환된 바이오패스웨이를 저장하고, 상기 노드 또는 관계와 관련된 속성정보를 추출하는 단계; 및 복수의 상기 저장된 바이오패스웨이를 통합하는 단계를 포함하는 패스웨이 통합 방법을 포함한다.

[0006] 또한 본 발명의 제 3 측면에 따르면, 적어도 하나 이상의 바이오패스웨이 입력을 받고, 상기 적어도 하나 이상의 바이오패스웨이를 제1언어로 변환하고, 상기 제1언어로 변환된 바이오패스웨이를 노드 또는 관계로 분리하고, 상기 노드 또는 관계를 표준용어로 표준화하고 상기 노드 또는 관계의 형태로 상기 제1언어로 변환된 바이오패스웨이를 저장하고, 상기 노드 또는 관계와 관련된 속성정보를 추출하고, 복수의 상기 저장된 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체를 포함한다.

발명의 효과

[0007] 이상에서 설명한 바와 같이 본 발명에 의하면 바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체를 제공함으로써 저비용으로 바이오패스웨이 데이터베이스를 효율적으로 구축할 수 있다.

[0008] 바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체를 제공함으로써 통합된 패스웨이를 통해 심층 지식 추출을 통한 신규 발명 기전 추출이 가능하다.

[0009] 또한 바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체를 제공함으로써 기술 발전에 맞춰 신속한 데이터베이스 확장 및 변경이 가능하다.

도면의 간단한 설명

[0010] 도 1은 본 발명의 일 실시예에 따른 바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체를 설명하기 위한 개략적인 흐름도,

도 2는 본 발명의 일 실시예에 따른 바이오패스웨이 통합을 위한 장치를 설명하기 위한 모듈 구성도,

도 3은 본 발명의 일 실시예에 따른 바이오패스웨이 통합의 전체적인 과정을 나타낸 도면,

도 4는 본 발명의 일 실시예에 따른 입력된 바이오패스웨이 정보와 저장된 바이오패스웨이 전체 정보를 나타낸 도면,

도 5는 본 발명의 일 실시예에 따른 입력된 또 다른 하나의 바이오패스웨이 정보와 저장된 바이오패스웨이 전체 정보를 나타낸 도면,

도 6은 본 발명의 일 실시예에 따른 통합된 바이오패스웨이 정보와 저장된 바이오패스웨이 전체 정보를 나타낸 도면,

도 7은 본 발명의 일 실시예에 따른 알츠하이머 병의 바이오패스웨이를 나타낸 도면,

도 8은 본 발명의 일 실시예에 따른 KGML언어로 표현된 바이오패스웨이를 나타낸 도면,

도 9는 본 발명의 일 실시예에 따른 바이오패스웨이 통합 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 본 발명의 일 실시예를 첨부된 도면들을 참조하여 상세히 설명한다. 또한, 본 발명을 설명함에 있어, 관련된 공

지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다.

- [0012] 도 1은 본 발명의 일 실시예에 따른 바이오패스웨이 통합을 위한 장치, 그 방법 및 바이오패스웨이들을 통합하는 프로그램을 저장하는 저장매체를 설명하기 위한 개략적인 흐름도이다.
- [0013] 도 1을 참조하면, 본 발명에 따른 바이오패스웨이 통합은 바이오패스웨이 입력(10), 하나의 언어로 통일하기 위한 언어통일부(11), 하나의 언어로 통일된 바이오패스웨이를 노드 또는 관계로 분리하여 표준용어로 표준화하여 저장하는 저장부(12), 통합을 위해 원하는 바이오패스웨이를 검색하는 검색부(13), 복수의 바이오패스웨이들을 통합하는 통합생성부(14)를 통해 진행된다.
- [0014] 도 2는 본 발명의 일 실시예에 따른 바이오패스웨이 통합을 위한 장치를 설명하기 위한 모듈 구성도이다.
- [0015] 도 2를 참조하면, 본 발명에 따른 바이오패스웨이 통합을 위한 장치는 입력모듈(20), 언어통일모듈(21), 분리모듈(22), 표준화모듈(23), 검색모듈(24), 및 통합모듈(25)을 포함한다. 입력모듈은(20)은 통합하기 원하는 바이오패스웨이 입력을 받는다.
- [0016] 언어통일모듈(21)은 다양한 기관에서 이질적인 언어로 표현된 바이오패스웨이를 하나의 언어로 통일한다. 이질적인 언어에는 KGML, SBML, BIOPAX, PSI-MI가 있다. 이질적인 언어로 표현된 바이오패스웨이를 하나의 언어로 통일시키기 위해서 SBML을 사용하지만 이는 사용자의 의도에 따라 변경하거나 새로운 언어 생성도 가능하다. 하나의 언어로 통일된 바이오패스웨이를 향후 통합 바이오패스웨이를 만드는데 사용하기 위해 바이오패스웨이 DB(26)에 저장할 수 있다.
- [0017] 분리모듈(22)은 바이오패스웨이를 노드 또는 관계로 분리한다. 노드의 예시로 단백질, 효소, 유전자 등이 있으며 관계는 노드와 노드 간의 메커니즘으로 예시로는 활성화, 발현, 전이 등이 있다. 구체적으로 바이오패스웨이를 노드-관계-노드 형태의 트리플형태로 변환하는 것이 바람직하지만 분리형태에서 관해서는 사용자의 의도에 따라 변경이 가능하며 트리플형태에 한정되지 않는다. 참고로 분리모듈은 바이오패스웨이로부터 분리되는 노드 또는 관계의 존재 여부 확인 등 전반적으로 바이오패스웨이로부터 노드 또는 관계를 분리하기 위한 필요한 정보를 얻기 위해 정보 DB(27)를 사용한다. 또한 분리된 바이오패스웨이를 향후 통합 바이오패스웨이를 만드는데 사용하기 위해 분리된 형태로 바이오패스웨이 DB(26)에 저장할 수 있다.
- [0018] 표준화모듈(23)은 하나의 언어로 통일된 바이오패스웨이로부터 분리된 노드 또는 관계가 실질적으로 동일한 기능을 지니더라도 다양한 용어로 표현될 수 있으므로 노드 또는 관계를 표준용어로 표준화한다. 참고로 의미가 동일하지만 다양한 용어로 표시된 노드란, 예를 들어 단백질의 한 종류인 ALDH1의 경우 Aldehyde dehydrogenase 1, Aldehyde dehydrogenase 1 family, member A1, ALHDII, ALDH-E1, Acetaldehyde dehydrogenase 1, Retinal dehydrogenase 1, RALDH 1, ALDC 등으로 다양하게 표현할 수 있다. 다른 표현으로 위와 같이 표현된 용어는 동일한 단백질을 나타내지만 사용자마다 다른 용어를 사용하는 경우가 많다. 참고로 표준화 모듈은 분리된 노드 또는 관계를 표준화하기 위하여 표준용어와 이와 의미가 동일하지만 표현이 다른 다양한 방언에 대한 매핑정보 등 분리된 노드 또는 관계를 표준화하기 위해 필요한 정보를 얻기 위해 정보DB(27)를 사용한다. 정보 DB의 예로 HPRD(Human Protein Reference Database)가 있지만 사용자의 의도에 따라 변경이 가능하므로 이에 한정하지 않는다. 또한 향후 통합 바이오패스웨이를 만드는데 사용하기 위해 표준화된 노드 또는 관계의 형태로 바이오패스웨이를 바이오패스웨이 DB(26)에 저장할 수 있다.
- [0019] 또한 표준화모듈(23)은 바이오패스웨이로부터 분리된 노드 또는 관계와 관련된 속성정보를 추출한다. 속성정보란 입력되는 바이오패스웨이의 노드별 빈도수, 노드별 링크수, 관계별 빈도수가 포함되며 또한 저장된 바이오패스웨이의 전체 노드별 빈도수, 전체 노드별 링크수, 전체 관계별 빈도수를 포함한다. 하지만 속성정보의 종류는 사용자의 의도에 따라서 여러 가지가 가능하므로 여기에 언급된 것으로 한정되지는 않는다. 참고로 표준모듈은 속성정보를 추출하는데 있어서 전반적으로 필요한 정보를 얻기 위해 바이오패스웨이 DB(26)나 정보 DB(27)를 사용한다. 또한 추출된 속성정보는 향후 통합 바이오패스웨이를 만드는데 사용하기 위해 바이오패스웨이 DB(26)에 저장할 수 있다.
- [0020] 검색모듈(24)은 통합하고자 하는 바이오패스웨이를 검색한다. 구체적인 검색방법은 노드, 관계, 트리플(노드-관계-노드), 바이오패스웨이 일부분, 바이오패스웨이 전체를 이용하여 완전일치 또는 유사도를 기반으로 검색한다. 하지만 검색방법은 사용자의 의도에 따라 변경이 가능하므로 이에 한정하지 않는다. 참고로 검색방법을 다양하게 하기 위해 정보 DB(27)를 활용할 수 있다. 통합모듈(25)은 검색모듈을 통해 검색된 바이오패스웨이 또는 사용자가 임의로 지정한 바이오패스웨이(28)를 통합한다. 통합모듈은 바이오패스웨이들을 트리플(노드-관

계-노드)를 기반으로 통합되지만 이에 한정하는 것은 아니며 사용자의 의도에 따라서 변경이 가능하다. 또한 통합된 바이오패스웨이와 통합된 바이오패스웨이의 노드 또는 관계와 관련된 속성정보는 향후 또 다른 통합 바이오패스웨이를 만드는 데 사용하기 위해 바이오패스웨이 DB(26)에 저장할 수 있다.

- [0021] 도 3은 본 발명의 일 실시예에 따른 바이오패스웨이 통합의 전체적인 과정을 나타낸 도면이다.
- [0022] 도 3를 참조하면, 본 발명에 따라 부호 30은 입력모듈(20)로 입력되는 바이오패스웨이를 의미하며 부호 31은 입력모듈로 입력되는 또 다른 하나의 바이오패스웨이를 의미한다. 부호 32 및 33은 입력된 바이오패스웨이들이 분리과정을 거쳐 노드-관계-노드, 일명 트리플형태로 분리된 것 중 공통된 트리플을 의미한다. 구체적으로 바이오패스웨이 통합 과정을 설명하면, 입력된 바이오패스웨이들(30,31)은 하나의 언어로 통일된 후 분리과정과 표준 용어로 표준화과정을 거쳐 공통된 트리플을 기반으로 통합된다. 도 3에서 알 수 있듯이, 입력된 바이오 패스웨이들(30,31)의 공통된 트리플은 'B-D'(32,33)이며 입력된 바이오패스웨이들은 공통된 트리플인 'B-D'(34)를 기반으로 통합된다.
- [0023] 도 4는 본 발명의 일 실시예에 따른 입력된 바이오패스웨이 정보와 저장된 바이오패스웨이 전체 정보를 나타낸 도면이다.
- [0024] 도 4를 참조하면, 본 발명에 따라 부호 40은 입력되는 바이오패스웨이이며 부호 41은 분리모듈을 통해 바이오패스웨이(40)가 노드 또는 관계에 기반하여 트리플형태로 분리된 것을 나타낸다. 부호 42는 도 3에서 전술한 바와 같이 공통된 트리플을 나타낸다. 부호 43은 입력되는 바이오패스웨이의 노드 또는 관계와 관련하여 추출된 속성 정보인 노드별 빈도수, 노드별 링크수를 나타내며 부호 44는 관계별 빈도수를 나타낸다. 속성정보들은 바이오패스웨이 DB에 저장되어 바이오패스웨이 DB에 이미 저장되어 있는 바이오패스웨이의 노드 또는 관계와 관련된 속성정보에 반영된다. 따라서 부호 45는 바이오패스웨이 DB에 저장되는 바이오패스웨이의 노드 또는 관계와 관련하여 추출된 속성정보인 전체 노드별 빈도수, 전체 노드별 링크수를 나타내며 부호 46은 전체 관계별 빈도수를 나타낸다. 구체적으로 설명하면, 입력되는 바이오패스웨이(40)의 노드 또는 관계와 관련된 속성정보를 살펴보면, 입력되는 바이오패스웨이(40)의 노드 A는 노드 B, C, D와 연결되어 링크수가 3개며 입력되는 바이오패스웨이 내에서 노드 A는 한 개만 존재하므로 빈도수는 1번이다. 입력되는 바이오패스웨이 내에서 트리플(관계)인 'B-D' 1개만 존재하므로 빈도수는 1번이다.
- [0025] 도 5는 본 발명의 일 실시예에 따른 입력된 또 다른 하나의 바이오패스웨이 정보와 저장된 바이오패스웨이 전체 정보를 나타낸 도면이다.
- [0026] 도 5를 참조하면, 본 발명에 따라 부호 50은 입력되는 바이오패스웨이이며 부호 51은 분리모듈을 통해 바이오패스웨이(50)가 노드 또는 관계에 기반하여 트리플형태로 분리된 것을 나타낸다. 부호 52는 도 3에서 전술한 바와 같이 공통된 트리플을 나타낸다. 부호 53은 입력되는 바이오패스웨이의 노드 또는 관계와 관련하여 추출된 속성 정보인 노드별 빈도수, 노드별 링크수를 나타내며 부호 54는 관계별 빈도수를 나타낸다. 속성정보들은 바이오패스웨이 DB에 저장되어 바이오패스웨이 DB에 이미 저장되어 있는 바이오패스웨이의 노드 또는 관계와 관련된 속성정보에 반영된다. 따라서 부호 55는 바이오패스웨이 DB에 저장되는 바이오패스웨이의 노드 또는 관계와 관련하여 추출된 속성정보인 전체 노드별 빈도수, 전체 노드별 링크수를 나타내며 부호 56은 전체 관계별 빈도수를 나타낸다. 구체적으로 설명하면, 입력되는 바이오패스웨이(50)의 노드 또는 관계와 관련된 속성정보를 살펴보면, 입력되는 바이오패스웨이(50)의 노드 B는 노드 D, E, F와 연결되어 링크수가 3개며 입력되는 바이오패스웨이 내에서 노드 B는 한 개만 존재하므로 빈도수는 1번이다. 입력되는 바이오패스웨이 내에서 트리플(관계)인 'B-D' 1 개만 존재하므로 빈도수는 1번이다. 또한 저장되어 있는 바이오패스웨이의 노드 또는 관계와 관련된 전체 속성정보를 살펴보면, 도 5에서는 또 다른 하나의 바이오패스웨이(50)가 추가로 입력되었기 때문에 도 5의 저장된 바이오패스웨이 전체 속성정보에서는 노드 B의 빈도수(55), 노드D의 빈도수(55), 트리플 'B-D' 빈도수(56)가 도 4의 저장된 바이오패스웨이 전체 속성정보와 비교하여 하나씩 증가했다.
- [0027] 도 6은 본 발명의 일 실시예에 따른 통합된 바이오패스웨이 정보와 저장된 바이오패스웨이 전체 정보를 나타낸 도면이다.
- [0028] 도 6을 참조하면, 본 발명에 따라 부호 63은 통합된 바이오패스웨이의 노드 또는 관계와 관련하여 추출된 속성 정보인 노드별 빈도수, 노드별 링크수를 나타내며 부호 65는 관계별 빈도수를 나타낸다. 속성정보들은 바이오패스웨이 DB에 저장되어 바이오패스웨이 DB에 이미 저장되어 있는 바이오패스웨이의 노드 또는 관계와 관련된 속성정보에 반영된다. 따라서 부호 67은 바이오패스웨이 DB에 저장되는 바이오패스웨이의 노드 또는 관계와 관련하여 추출된 속성정보인 전체 노드별 빈도수, 전체 노드별 링크수를 나타내며 부호 69는 전체 관계별 빈도수를

나타낸다. 구체적으로 설명하면, 통합된 바이오패스웨이의 노드 또는 관계와 관련된 속성정보를 살펴보면, 통합된 바이오패스웨이의 노드 B는 노드 A, D, E와 연결되어 링크수가 3개이다. 입력되는 바이오패스웨이 내에서 노드 D(62)는 도 4와 도 5에서 전술한 바와 같이 통합되기 전의 바이오패스웨이에서 각각 한 개씩 존재하였으므로 통합된 바이오패스웨이 내에서 빈도수는 2번이다. 또한 통합된 바이오패스웨이 내에서 트리플(관계)인 'B-D'(64)는 도 4와 도 5에서 전술한 바와 같이 통합되기 전의 바이오패스웨이에서 각각 한 개씩 존재하였으므로 통합된 바이오패스웨이 내에서 두 개가 존재하므로 빈도수는 2번이다. 또한 저장되어 있는 바이오패스웨이의 노드 또는 관계와 관련된 전체 속성정보를 살펴보면, 도 6에서 노드 D의 빈도수(66)가 다른 노드에 비해 현저히 낮다. 이는 곧 특정 바이오패스웨이에서만 집중적으로 나타나는 것을 의미하며 모든 바이오패스웨이에 비해 중요도가 높은 핵심정보라고 할 수 있다. 따라서 중요성을 나타내기 위해 노드 D를 다른 노드에 비해 확대(61)하여 표현한다. 또한 트리플 'B-D'(68)의 경우 다른 트리플 즉 다른 관계정보에 비해 빈도수가 현저히 낮음을 확인할 수 있다. 이는 곧 특정 바이오패스웨이에서만 집중적으로 나타나는 것을 의미하며 모든 바이오패스웨이에 비해 중요도가 높은 주요 매커니즘 즉, 핵심정보라고 할 수 있다. 따라서 중요성을 나타내기 위해 트리플 'B-D' 다른 트리플에 비해 굵게 표시(60)한다. 표현방식은 사용자의 의도에 따라 변경이 가능하므로 이에 한정하지 않는다.

- [0029] 도 7은 본 발명의 일 실시예에 따른 알츠하이머 병의 바이오패스웨이를 나타낸 도면이다.
- [0030] 도 7을 참조하면, 본 발명에 따라 부호 70은 바이오패스웨이 내에서 사각형으로 표현된 노드를 의미하며 부호 71은 바이오패스웨이 내에서 화살표로 표현된 관계를 의미한다.
- [0031] 도 8은 본 발명의 일 실시예에 따른 KGML언어로 표현된 바이오패스웨이를 나타낸 도면이다.
- [0032] 도 8을 참조하면, 본 발명에 따라 도 8은 도 7의 일부분을 확대하여 앞에서 전술한 KGML언어로 바이오패스웨이를 표현한 것이다. 부호 80은 'FADD->CASP80' 트리플을 나타내며 이는 XML로 텍스트 표현(81)이 가능하다. 구체적으로 설명하면, <graphics name>의 뒤에 나오는 내용이 노드명(CASP8,82)을 의미하고 CASP8, CASP-8 등 동일한 용어이지만 다른 형태의 용어들이 쉼표(,)로 분리되어 나올 수 있다. 이 노드명들은 정보 DB를 이용하여 표준화되고, 정보 DB(27) 확장에도 사용될 수 있다.
- [0033] 또한 <relation>의 뒤에 나오는 내용은 관계정보(PPRe1, 83)를 의미한다. 예를 들어 PPRe1은 Protein Protein Relation의 약어로 단백질-단백질 관계를 의미하며, PCRe1은 Protein Compound Relation의 약어로 단백질-복합물 관계를 의미한다.
- [0034] 도 9는 본 발명의 일 실시예에 따른 바이오패스웨이 통합 방법을 설명하기 위한 흐름도이다.
- [0035] 도 9를 참조하면, 본 발명에 따라 바이오패스웨이 통합 방법은 바이오패스웨이 입력을 받는 단계(S90), 바이오패스웨이를 제1언어로 변환하는 단계(S91), 바이오패스웨이를 분리하는 단계(S92), 바이오패스웨이의 노드 또는 관계를 표준용어로 표준화하고 노드 또는 관계 형태로 바이오패스웨이를 저장하고, 노드 또는 관계와 관련된 속성정보를 추출하는 단계(S93), 바이오패스웨이를 검색하는 단계(S94), 복수의 바이오패스웨이들을 통합하는 단계(S95)를 포함한다.
- [0036] 입력 모듈(20)은 바이오패스웨이 입력을 받는 단계(S90)로 변환될 수 있으며 이에 대한 구체적인 설명은 도 2에서 전술한 바 있다.
- [0037] 언어통일모듈(21)은 바이오패스웨이를 제1언어로 변환하는 단계(S91)로 변환될 수 있고 이에 대한 구체적인 설명은 도 2에서 전술한 바 있다.
- [0038] 분리모듈(22)은 바이오패스웨이를 분리하는 단계(S92)로 변환될 수 있으며 이에 대한 구체적인 설명은 도 2, 도 3, 도 4, 도 5 및 도 6에서 전술한바 있다.
- [0039] 표준화모듈(23)은 바이오패스웨이의 노드 또는 관계를 표준용어로 표준화하고 노드 또는 관계 형태로 바이오패스웨이를 저장하고, 노드 또는 관계와 관련된 속성정보를 추출하는 단계(S93)로 변환될 수 있으며 이에 대한 구체적인 설명은 도 2, 도 3, 도 4, 도 5 및 도 6에서 전술한 바 있다.
- [0040] 검색모듈(24)은 바이오패스웨이를 검색하는 단계(S94)로 변환될 수 있으며 이에 대한 구체적인 설명은 도 2에서 전술한 바 있다.
- [0041] 통합모듈(25)은 복수의 바이오패스웨이들을 통합하는 단계(S95)로 변환될 수 있으며 이에 대한 구체적인 설명은 도 2, 도 3, 도 4, 도 5 및 도 6에서 전술한바 있다.

[0042]

본 발명의 명세서에 개시된 실시예들은 본 발명을 한정하는 것이 아니다. 본 발명의 범위는 아래의 특허청구범위에 의해 해석되어야 하며, 그와 균등한 범위 내에 있는 모든 기술도 본 발명의 범위에 포함되는 것으로 해석해야 할 것이다.

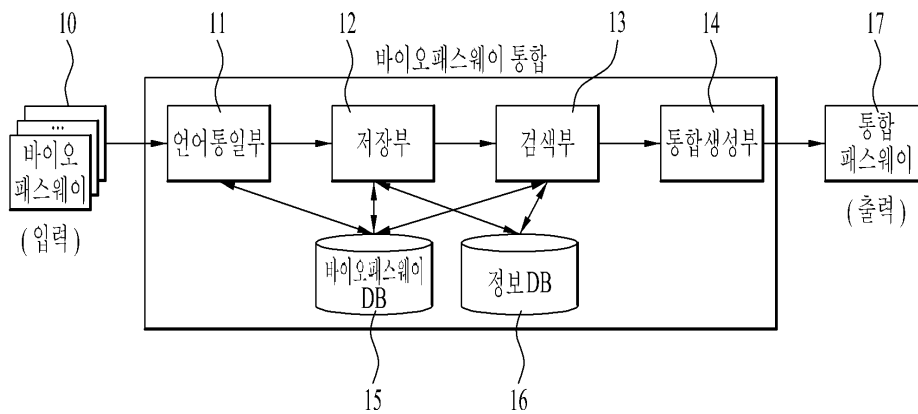
부호의 설명

[0043]

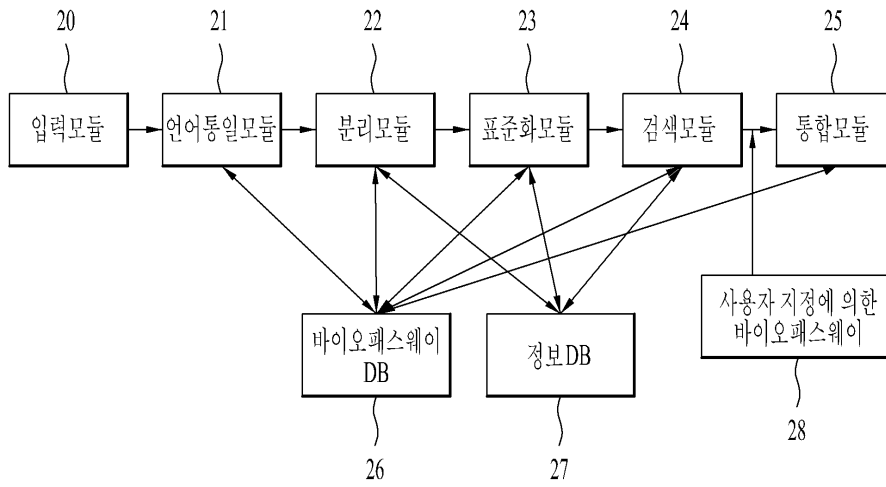
- 20 : 입력모듈
- 21 : 언어통일모듈
- 22 : 분리모듈
- 23 : 표준화모듈
- 24 : 검색모듈
- 25 : 통합모듈
- 26 : 바이오패스웨이 DB
- 27 : 정보 DB

도면

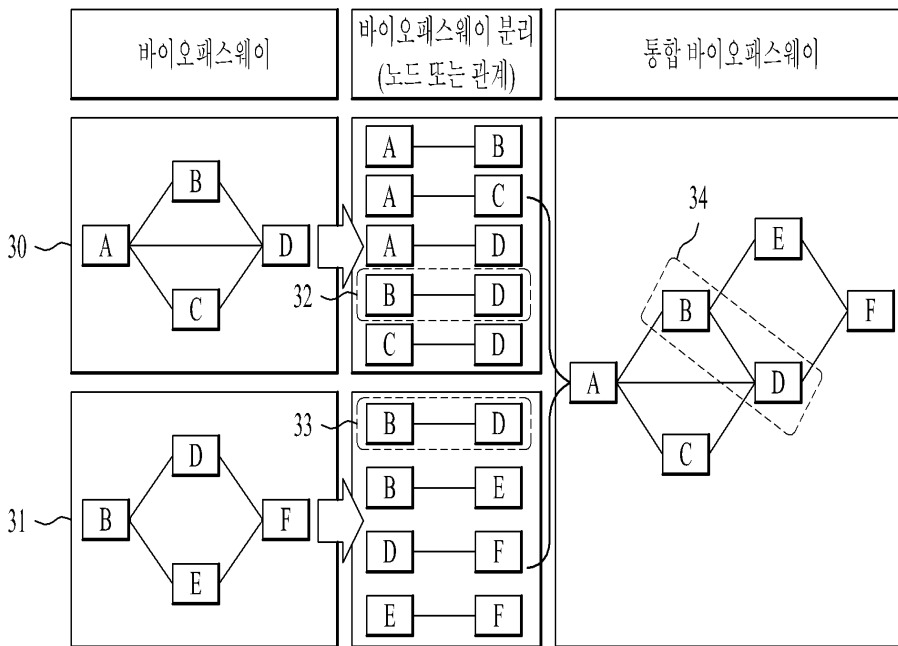
도면1



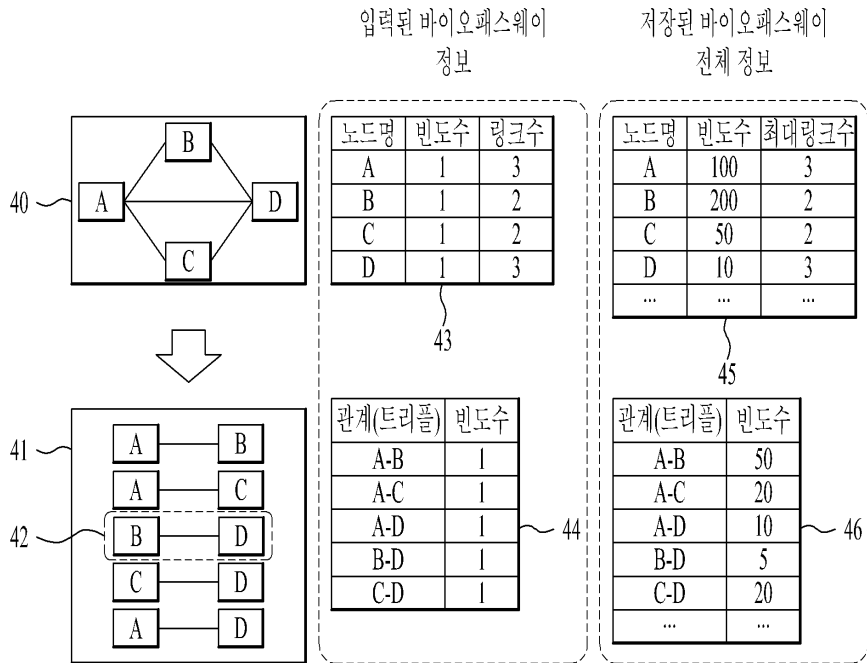
도면2



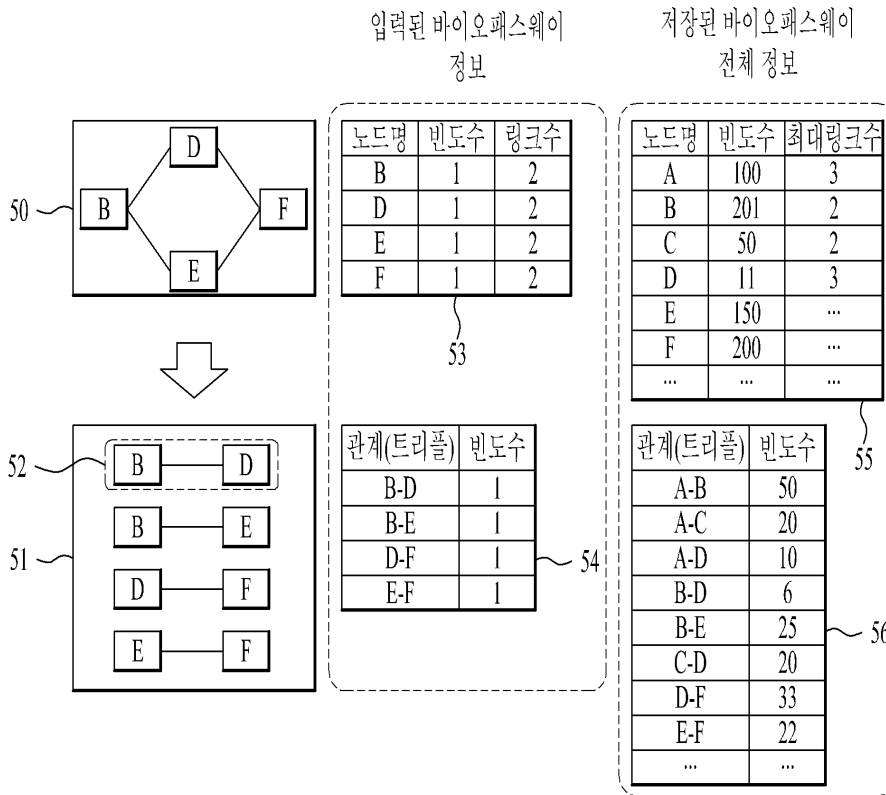
도면3



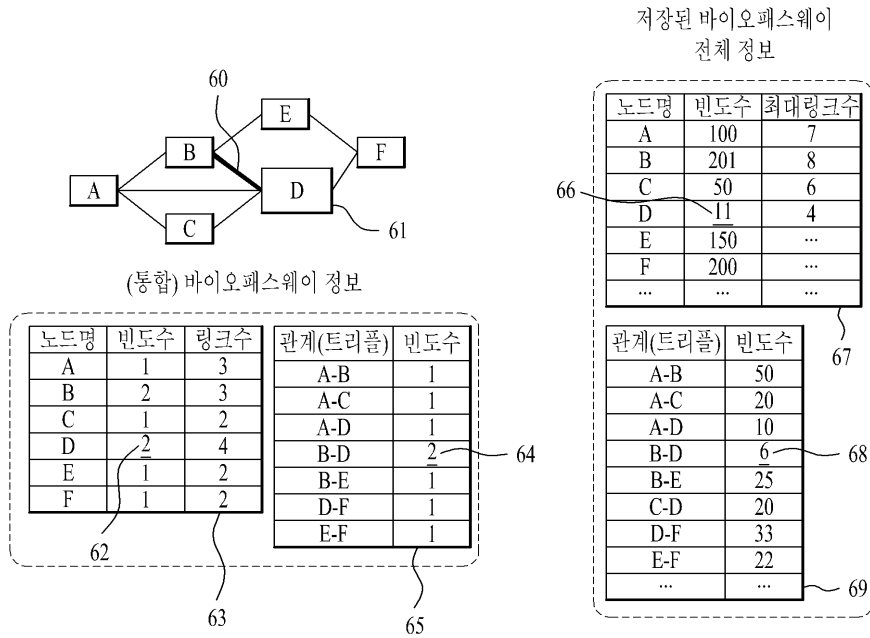
도면4



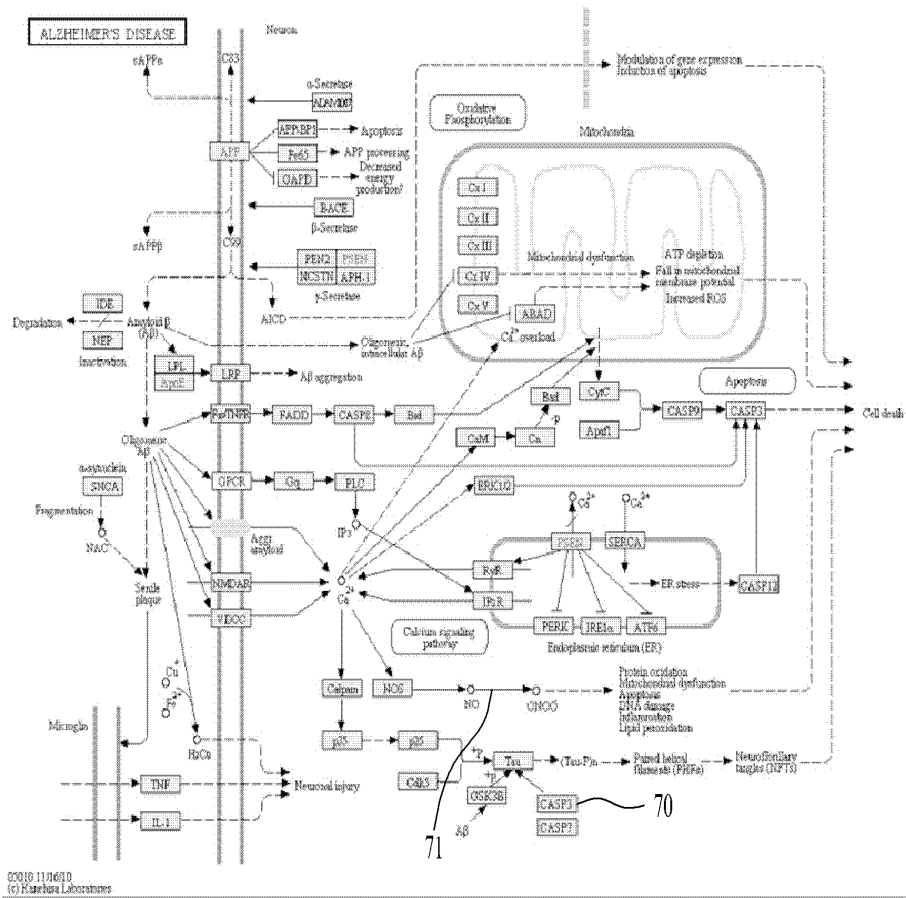
도면5



도면6

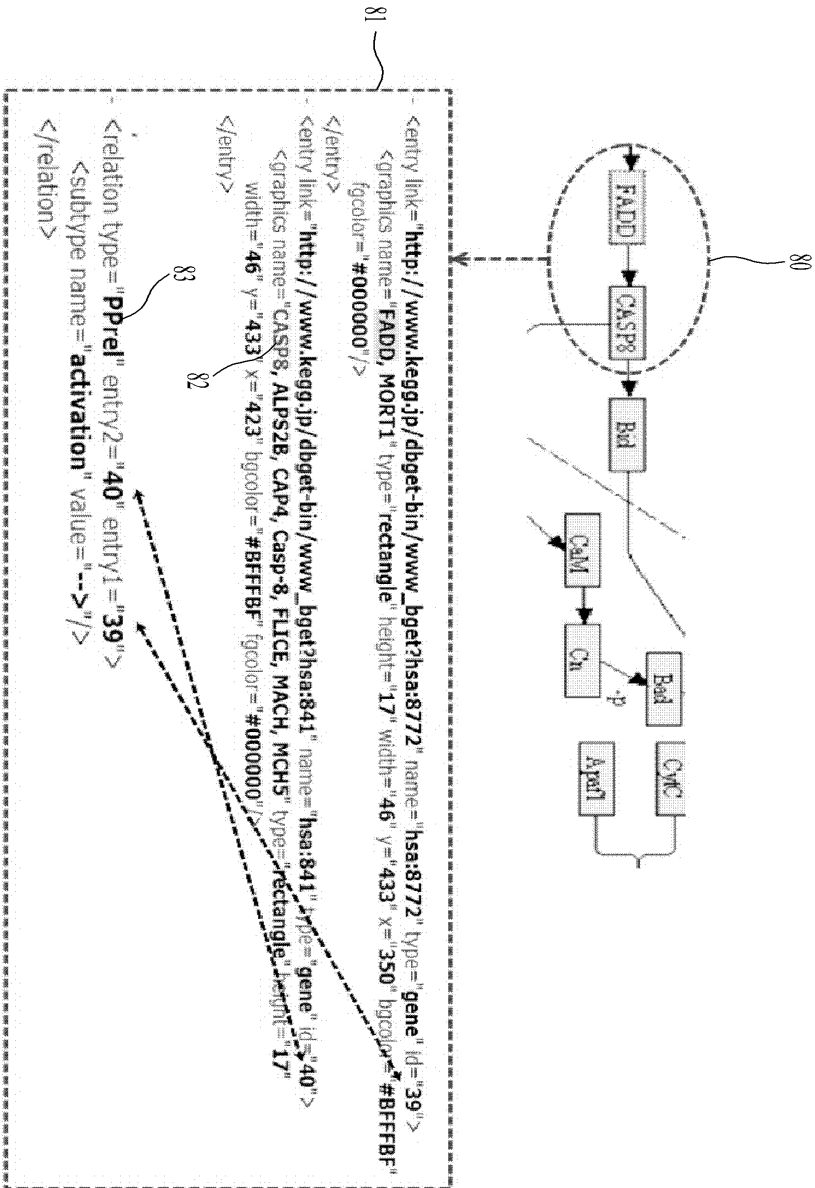


도면7



02010 117610
(c) Hanhara Laboratories

도면8



도면9

