



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2014년10월21일
 (11) 등록번호 10-1448731
 (24) 등록일자 2014년10월01일

(51) 국제특허분류(Int. Cl.)
 G06F 19/00 (2011.01) G06F 17/30 (2006.01)
 (21) 출원번호 10-2010-0105299
 (22) 출원일자 2010년10월27일
 심사청구일자 2013년02월22일
 (65) 공개번호 10-2012-0043977
 (43) 공개일자 2012년05월07일
 (56) 선행기술조사문헌
 Tae-Kyung Kim 외 5인, 'PubMine: An Ontology-Based Text Mining System for Deducing Relationships among Biological Entities', Interdisciplinary Bio Central, 2011.04.25.

(73) 특허권자
 한국생명공학연구원
 대전광역시 유성구 과학로 125 (어은동)
 (72) 발명자
 김태경
 대전광역시 유성구 과학로 125 (어은동)
 오정수
 대전광역시 유성구 과학로 125 (어은동)
 (뒷면에 계속)
 (74) 대리인
 이원희

전체 청구항 수 : 총 13 항

심사관 : 박재용

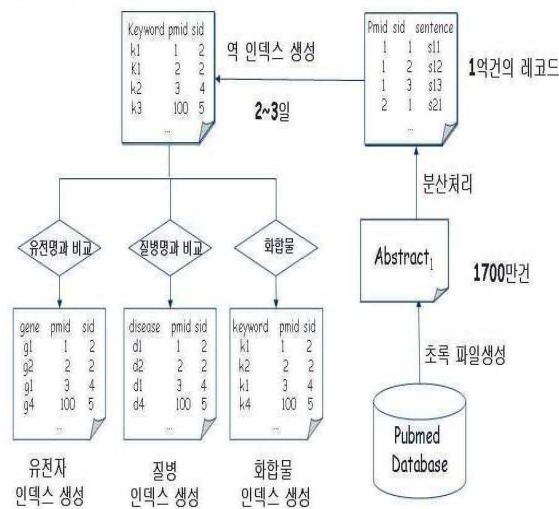
(54) 발명의 명칭 **다차원 인덱스를 이용하여 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법**

(57) 요약

본 발명은 유전자-질병-화합물(gene-disease-chemical)에 대한 다차원 분석을 위해 대용량의 생명공학 문헌으로부터 다차원 인덱스를 적용하여 유전자-질병-화합물 관계(gene-disease-chemical relationship)를 효과적으로 추출함으로써, 검색의 효율성 및 정확도를 높이고 유전자-질병-화합물에 대한 다차원 분석을 가능하게 하기 위한 방법에 관한 것으로, 본 발명에 따르면, 다차원 인덱스 구조를 통해 유전자-질병-화합물간의 관계를 신속하고 유연하게 추출할 수 있는 동시에, 유전자-질병-화합물의 관계를 문장수준에서 검색하고 확인할 수 있도록 하며, 인덱스를 활용하여 직관적인 사용자 인터페이스를 구현할 수 있는 다차원 인덱스를 이용하여 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법이 제공된다.

대표도 - 도1

대용량 문헌으로부터 질병-유전자-화합물에 대한 다차원 인덱스 구축절차



(72) 발명자

이상혁

대전광역시 유성구 과학로 125 (어은동)

허보경

대전광역시 유성구 과학로 125 (어은동)

이 발명을 지원한 국가연구개발사업

과제고유번호 20100002065

부처명 교육과학기술부

연구관리전문기관 한국연구재단

연구사업명 미래기반기술개발사업

연구과제명 생명정보 연구성과물 등록 활용 사업

기 여 율 1/1

주관기관 한국생명공학연구원

연구기간 2010.04.01 ~ 2011.03.31

특허청구의 범위

청구항 1

인덱스를 생성하는 수단, 상기 인덱스를 관점별로 저장하는 수단 및 상기 인덱스의 관계를 분석하는 수단을 구비하는 컴퓨터가 수행하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서,

상기 인덱스를 생성하는 수단이 대용량 생명공학 문헌을 입력받아 질병, 유전자, 화합물에 대한 다차원 인덱스를 구축하는 단계와,

상기 인덱스를 관점별로 저장하는 수단이 구축된 상기 다차원 인덱스를 미리 정해진 인덱스 저장구조에 따라 저장하는 단계와,

상기 인덱스의 관계를 분석하는 수단이 상기 저장된 인덱스를 기초로 사용자에게 의해 입력된 검색어를 이용하여 상기 대용량 생명공학 문헌으로부터 질병, 유전자, 화합물에 대한 다차원 분석을 수행하는 단계를 포함하여 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 2

제 1항에 있어서,

상기 다차원 인덱스를 구축하는 단계는,

PubMed 데이터베이스로부터 초록(abstract) 필드가 널(NULL)이 아닌 문헌만을 추출하는 단계와,

각 초록의 내용을 문장 단위로 분할하여 큐레이션(curation)을 통해 문장 테이블(sentence table)로 만드는 단계와,

상기 문장 테이블에 대하여 역인덱스를 구축하는 단계와,

유전자, 질병, 화합물에 대한 각각의 동의어 사전을 상기 역인덱스와 비교하여 유전자, 질병, 화합물에 대한 각각의 차원 인덱스를 구축하는 단계를 포함하여 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 3

제 2항에 있어서,

상기 문장 테이블로 만드는 단계에서,

상기 문장 테이블은, [pubmed id, 문장 id(sentence id), 문장(sentence)]의 순서로 저장되는 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 4

제 1항에 있어서,

상기 인덱스 저장구조는, 스타 스키마(Star Schema) 구조인 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 5

제 4항에 있어서,

상기 인덱스 저장구조에서, 질병 인덱스는, pubmed 아이디, 문장번호, 질병 아이디 및 질병이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 표준 질병명 및 질병에 대한 동의어 정보가 관련지어져 저장되는 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 6

제 4항에 있어서,

상기 인덱스 저장구조에서, 유전자 인덱스는, pubmed 아이디, 문장번호, 유전자 아이디 및 유전자 이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 표준 유전자 및 유전자에 대한 동의어 정보가 관련지어져 저장되는 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 7

제 4항에 있어서,

상기 인덱스 저장구조에서, 화합물 인덱스는, pubmed 아이디, 문장번호, 합성물 아이디 및 화합물 아이디, 화합물 이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 합성물명 및 합성물에 대한 동의어 정보가 관련지어져 저장되는 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 8

제 5항 내지 제 7항 중 어느 한 항에 있어서,

상기 인덱스 저장구조는, 질병 인덱스, 유전자인덱스 및 화합물 인덱스 이외에 다른 분석차원에 대하여도 인덱스 정보를 추가하여 다차원 분석모델을 수립할 수 있도록 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 9

제 1항에 있어서,

상기 방법은,

상기 대용량 생명공학 문헌을 1차원(유전자, 질병, 화합물), 2차원(유전자-질병, 질병-유전자, 유전자-화합물, 화합물-유전자, 질병-화합물, 화합물-질병 관계) 및 3차원(유전자-질병-화합물 관계)으로 분석할 수 있도록 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 10

제 1항에 있어서,

사용자가 검색어를 입력하면, 검색결과를 나타내는 화면에 각 유전자, 질병, 화합물에 대한 색 또는 하이라이트를 각각 적용하여, 시각적인 효과를 줄 뿐만 아니라 사용자가 직관적으로 내용을 이해할 수 있도록 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 11

제 1항에 있어서,

사용자가 검색어를 입력하면, 문장을 기준으로 내용을 파악한 다음 전체 초록을 볼 수 있도록 함으로써, 문장을 중심으로 초록 내용을 확인할 수 있도록 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 12

제 1항에 있어서,

사용자가 검색어를 입력하면, 그와 관련된 키워드들이 표시되고, 상기 키워드를 선택하면 상기 검색어 및 상기 키워드에 해당하는 검색결과 및 초록이 표시되어, 사용자가 상기 검색어와 상기 키워드간의 관계 분석을 용이하게 수행할 수 있도록 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법.

청구항 13

제 1항에 있어서,

유전자-질병-화합물간의 관계를 추출하기 위한 별도의 프로그램을 작성할 필요 없이, SQL을 이용하여 상기 인덱스에 접근하여 필요한 분석을 즉시 수행할 수 있도록 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터

유전자-질병-화합물 관계를 추출하는 방법.

명세서

기술분야

[0001] 본 발명은, 생명정보학분야의 텍스트마이닝(text mining) 기술에 관한 것으로, 더 상세하게는, 유전자-질병-화합물(gene-disease-chemical)에 대한 다차원 분석을 위해 대용량의 생명공학 문헌으로부터 다차원 인덱스를 적용하여 유전자-질병-화합물 관계(gene-disease-chemical relationship)를 효과적으로 추출함으로써, 검색의 효율성 및 정확도를 높이고 유전자-질병-화합물에 대한 다차원 분석을 가능하게 하기 위한 방법에 관한 것이다.

배경기술

[0002] 종래, 생물학 분야의 연구에 있어서, 해마다 대량의 생물학 실험결과가 문헌으로 공개되고 있으며, 그에 따라 이러한 정보의 전략적 활용이 점점 중요해지고 있다.

[0003] 또한, 현재 바이오 텍스트 문헌으로부터 유전자-질병-화합물 관계를 파악하기 위해서는 PubMed 상에서 키워드 검색을 통해 확인하는 방법이 유일한 방법이나, PubMed 상에는 현재도 약 1만여 건의 문헌이 관리되고 있으며, 앞으로 그 양은 급격히 증가할 것이 확실시되고 있다.

[0004] 따라서 이와 같은 대용량 문헌으로부터 관심 있는 정보를 신속하게 확인하여 생명현상에 대한 검증, 확인 및 유추를 가능하게 하는 인프라에 대한 요구가 날로 증가되고 있다.

[0005] 상기한 바와 같이 바이오 텍스트 문헌으로부터 유전자-질병-화합물 관계를 파악하기 위한 종래기술의 예로서는, 예를 들면, 2008년 5월 16일자로 발간된 "PolySearch : aweb-based textmining system for extracting relationships between human diseases,genes,mutations,drugs and metabolites", Nucleic Acids Research, 2008. Vol.36과 같은 것이 있다.

[0006] 즉, 상기 문헌은, 쿼리(query)를 이용하여 해당 질병 또는 유전자를 통해 관련된 돌연변이 증상 및 약품 등을 검색할 수 있도록 하는 시스템을 개시하고 있다.

[0007] 그러나 상기 문헌에 개시된 유전자-질병-화합물 관계분석 기법은, X->Y의 관계만 고려할 뿐, X,Y->Z의 분석은 할 수 없다는 단점이 있는 것이었다.

[0008] 또한, 상기한 바와 같은 종래기술의 다른 예로서, 예를 들면, 2005년 2월 22일자로 발행된 "Integration of text- and data-mining using ontologies successfully selects disease gene candidates", Nucleic Acids Research, 2005. Vol.33, No.5에는, 온톨로지 기술과 텍스트마이닝, 데이터마이닝 기술을 이용하여 질병을 일으키는 유전자 후보를 선택하는 것에 대한 기술내용이 기재되어 있다.

[0009] 아울러, 상기한 바와 같은 종래기술의 다른 예로서, 예를 들면, 2005년 6월 28일자로 발행된 "Text-mining and information-retrieval services for molecular biology", Genome Biology, 2005. 6:224(10.1186/gb-2005-6-7-224)에는, 분자생물학에서 텍스트마이닝을 통해 유전자와 단백질간의 기능적 관계를 텍스트로부터 자동적으로 추출하는 기술내용이 개시되어 있다.

[0010] 그러나 상기한 바와 같은 종래기술에 기재된 방법들을 이용하여 대용량 생명공학 문헌으로부터 키워드 기반의 검색으로 유전자-질병-화합물 관계를 파악하는데에는 다음과 같은 한계가 있는 것이었다.

[0011] 첫째, 상기한 바와 같은 종래기술의 방법들은 질의의 대상이 초록이므로, 실제로는 음성인데 결과는 양성이라고 나오는 이른바 거짓 양성(false positive)이 증가하여, 검색되는 문헌의 양이 필요 이상으로 방대해지고, 그에 따라 사용자가 해당 정보를 확인하는데 시간이 많이 걸리게 되는 문제가 있었다.

[0012] 둘째로, 상기한 바와 같은 종래기술의 방법들은 유전자, 질병, 화합물에 대한 하이라이트 기능이 없어 사용자가 해당 문장을 한눈에 용이하게 식별하기 어려운 단점도 있었다.

[0013] 셋째로, 상기한 바와 같은 종래기술의 방법들은 유전자-질병-화합물간의 관계에 대한 요약정보를 제시하지 못하는 것이 대부분이며, 요약정보를 제시해주는 경우도 대부분 수작업의 결과로써 새로운 정보를 실시간으로 수용하지 못하는 한계가 있었다.

[0014] 따라서 상기한 바와 같은 종래기술의 문제점들을 해결하여, 유전자-질병-화합물간의 관계를 신속하고 유연하게 추출할 수 있고, 유전자-질병-화합물의 관계를 문장수준에서 검색하고 확인할 수 있도록 하는 동시에, 인덱스를 활용하여 직관적인 사용자 인터페이스 구현할 수 있는 새로운 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계 추출방법을 제공하는 것이 바람직하나, 아직까지 그러한 요구를 모두 만족시키는 시스템이나 방법은 제공되지 못하고 있는 실정이다.

발명의 내용

해결하려는 과제

[0015] 본 발명은 상기한 바와 같은 종래기술의 문제점들을 해결하기 위한 것으로, 따라서 본 발명의 목적은, 다차원 인덱스 구조를 통해 유전자-질병-화합물간의 관계를 신속하고 유연하게 추출할 수 있는 동시에, 유전자-질병-화합물의 관계를 문장수준에서 검색하고 확인할 수 있도록 하며, 인덱스를 활용하여 직관적인 사용자 인터페이스를 구현할 수 있는 다차원 인덱스를 이용하여 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법을 제공하고자 하는 것이다.

과제의 해결 수단

[0016] 상기한 바와 같은 목적을 달성하기 위해, 본 발명에 따르면, 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서, 상기 대용량 생명공학 문헌으로부터 질병, 유전자, 화합물에 대한 다차원 인덱스를 구축하는 단계와, 구축된 상기 다차원 인덱스를 미리 정해진 인덱스 저장구조에 따라 저장하는 단계와, 저장된 상기 인덱스를 이용하여, 사용자가 검색어를 입력하고 상기 대용량 생명공학 문헌으로부터 질병, 유전자, 화합물에 대한 다차원 분석을 수행하는 단계를 포함하여 구성된 것을 특징으로 하는 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법이 제공된다.

[0017] 여기서, 상기 다차원 인덱스를 구축하는 단계는, PubMed 데이터베이스로부터 초록(abstract) 필드가 널(NULL)이 아닌 문헌만을 추출하는 단계와, 각 초록의 내용을 문장 단위로 분할하여 큐레이션(curation)을 통해 문장 테이블(sentence table)로 만드는 단계와, 상기 문장 테이블에 대하여 역인덱스를 구축하는 단계와, 유전자, 질병, 화합물에 대한 각각의 동의어 사전을 상기 역인덱스와 비교하여 유전자, 질병, 화합물에 대한 각각의 차원 인덱스를 구축하는 단계를 포함하여 구성된 것을 특징으로 한다.

[0018] 또한, 상기 문장 테이블로 만드는 단계에서, 상기 문장 테이블은, [pubmed id, 문장 id(sentence id), 문장(sentence)]의 순서로 저장되는 것을 특징으로 한다.

[0019] 아울러, 상기 인덱스 저장구조는, 스타 스키마(Star Schema) 구조인 것을 특징으로 한다.

[0020] 또한, 상기 인덱스 저장구조에서, 질병 인덱스는, pubmed 아이디, 문장번호, 질병 아이디 및 질병이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 표준 질병명 및 질병에 대한 동의어 정보가 관련지어져 저장되는 것을 특징으로 한다.

[0021] 더욱이, 상기 인덱스 저장구조에서, 유전자 인덱스는, pubmed 아이디, 문장번호, 유전자 아이디 및 유전자 이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 표준 유전자 및 유전자에 대한 동의어 정보가 관련지어져 저장되는 것을 특징으로 한다.

[0022] 아울러, 상기 인덱스 저장구조에서, 화합물 인덱스는, pubmed 아이디, 문장번호, 합성물 아이디 및 화합물 아이디, 화합물 이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 합성물명 및 합성물에 대한 동의어 정보가 관련지어져 저장되는 것을 특징으로 한다.

[0023] 또한, 상기 인덱스 저장구조는, 상기 질병 인덱스, 상기 유전자인덱스 및 상기 화합물 인덱스 이외에 다른 분석 차원에 대하여도 인덱스 정보를 추가하여 다차원 분석모델을 수립할 수 있도록 구성된 것을 특징으로 한다.

[0024] 또, 상기 방법은, 상기 문장을 1차원(유전자, 질병, 화합물), 2차원(유전자-질병, 질병-유전자, 유전자-화합물, 화합물-유전자, 질병-화합물, 화합물-질병 관계) 및 3차원(유전자-질병-화합물 관계)으로 분석할 수 있도록 구성된 것을 특징으로 한다.

[0025] 또한, 상기 방법은, 사용자가 검색어를 입력하면, 검색결과를 나타내는 화면에 각 유전자, 질병, 화합물에 대한 색 또는 하이라이트를 각각 적용하여, 시각적인 효과를 줄 뿐만 아니라 사용자가 직관적으로 내용을 이해할 수

있도록 구성된 것을 특징으로 한다.

- [0026] 아울러, 상기 방법은, 사용자가 검색어를 입력하면, 문장을 기준으로 내용을 파악한 다음 전체 초록을 볼 수 있도록 함으로써, 문장을 중심으로 초록 내용을 확인할 수 있도록 구성된 것을 특징으로 한다.
- [0027] 더욱이, 상기 방법은, 사용자가 검색어를 입력하면, 그와 관련된 키워드들이 표시되고, 상기 키워드를 선택하면 상기 검색어 및 상기 키워드에 해당하는 검색결과 및 초록이 표시되어, 사용자가 상기 검색어와 상기 키워드간의 관계 분석을 용이하게 수행할 수 있도록 구성된 것을 특징으로 한다.
- [0028] 또한, 상기 방법은, 유전자-질병-화합물간의 관계를 추출하기 위한 별도의 프로그램을 작성할 필요 없이, SQL을 이용하여 상기 인덱스에 접근하여 필요한 분석을 즉시 수행할 수 있도록 구성된 것을 특징으로 한다.

발명의 효과

- [0029] 상기한 바와 같이, 본 발명에 따르면, 역인덱스, 다차원 인덱스를 활용하여 유전자-질병-화합물간의 관계를 신속하게 추출할 수 있도록 구성되어, 문장 단위의 정교한 검색을 지원하고, X->Y의 분석뿐만 아니라 X,Y->Z의 관계도 추출할 수 있는 다차원 인덱스를 이용하여 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법이 제공된다.
- [0030] 즉, 본 발명에 따르면, PubMed 데이터베이스에서 초록파일을 가져오고, 각 초록을 문장 단위로 분리하며, 분리된 문장에 있는 유전자, 질병, 화합물의 위치에 대한 역인덱스를 생성한 후, 유전자, 질병, 화합물에 각각에 대해 이름과 함께 차원 인덱스를 생성하며, 인덱스 생성시 동의어 용어사전을 활용하여 검색 정확도를 높이고, 인덱스와 문장을 연결하여 다차원 분석이 가능하도록 구성된 다차원 인덱스를 이용하여 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법이 제공된다.
- [0031] 따라서 본 발명에 따르면, 대용량 문헌으로부터 생명공학 개체들 간의 관계를 도출할 수 있으며, 아울러 이는, 생명공학뿐만 아니라 화학 및 물리학 등 다양한 분야의 문헌으로부터 새로운 관계정보를 도출하는 데에도 적용할 수 있다.

도면의 간단한 설명

- [0032] 도 1은 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서, 대용량 문헌으로부터 질병-유전자-화합물에 대한 다차원 인덱스를 구축하는 절차를 설명하기 위한 도면이다.
- 도 2는 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서, 질병-유전자-화합물의 관계를 추출하기 위한 저장 구조를 설명하기 위한 도면이다.
- 도 3은 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서, 다차원 분석구조를 적용하여 추출한 기본 검색결과를 나타내는 화면을 나타내는 도면이다.
- 도 4는 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서, 추출된 문장의 전체 초록내용을 제공하는 화면을 나타내는 도면이다.
- 도 5는 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서, 다차원 분석을 위한 입력화면과 결과화면을 나타내는 도면이다.
- 도 6은 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 있어서, 화합물-질병간의 관계를 추출하기 위한 SQL의 구조를 나타내는 도면이다.
- 도 7은 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 대한 실시예로서, 유전자-화합물 관계 분석화면을 나타내는 도면이다.
- 도 8은 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 대한 실시예로서, 유전자-유전자 관계 분석화면을 나타내는 도면이다.
- 도 9는 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 대한 실시예로서, 질병-유전자-초록 관계 분석화면을 나타내는 도면이다.
- 도 10은, 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에 대한 실시예로서, 질병-유전자 관계 분석화면을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0033] 이하, 첨부된 도면을 참조하여, 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법의 상세한 내용에 대하여 설명한다.
- [0034] 여기서, 이하에 설명하는 내용은 본 발명을 실시하기 위한 하나의 실시예일 뿐이며, 본 발명은 이하에 설명하는 실시예의 내용으로만 한정되는 것은 아니라는 사실에 유념해야 한다.
- [0035] 즉, 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법은, 후술하는 바와 같이 하여, 대용량 문헌으로부터 스타 스키마(Star Schema) 형태의 유전자-질병-화합물 분석을 위한 다차원 인덱스 구조 및 그러한 인덱스를 활용하여 검색결과에 포함된 유전자-질병-화합물에 대한 하이라이트 기능을 가지는 다차원 인덱스를 이용하는 유전자-질병-화합물 관계분석기법에 관한 것이다.
- [0036] 또한, 본 발명은, 예를 들면, 생명공학자들이 관심이 있는 질병과 관련하여 최근에 밝혀진 유전자 검색이나, 개체(Organism), 신체부위(Anatomy) 등의 차원을 추가하여 다양한 검색서비스 지원 및 생명공학 전 분야에서 사용할 수 있는 정보검색 시스템 등에 적용할 수 있다.
- [0037] 계속해서, 도 1 내지 도 10을 참조하여, 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법의 구체적인 구성에 대하여 설명한다.
- [0038] 먼저, 도 1을 참조하면, 도 1은 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법에서 다차원 인덱스의 구축과정을 나타내고 있다.
- [0039] 즉, 도 1에 나타난 바와 같이, 대용량 문헌으로부터 질병-유전자-화합물에 대한 다차원 인덱스를 구축하는 절차는, 먼저, PubMed 데이터베이스에서 초록파일을 가져온 후, 각 초록을 문장단위로 분리하고, 문장에 들어있는 유전자, 질병, 화합물에 대하여 각각의 위치에 대한 인덱스를 생성한다.
- [0040] 여기서, 상기한 각각의 인덱스 구축시, 동의어 용어사전을 활용하여 검색 정확도를 높이도록 구성한다.
- [0041] 상기한 바와 같이 인덱스를 생성한 후, 각각의 인덱스와 문장을 연결하여 사용자가 다차원 분석을 할 수 있도록 한다.
- [0042] 더 상세하게는, 상기한 다차원 인덱스를 구축하는 절차는, 먼저, PubMed 데이터베이스로부터 전체 문헌을 추출하며, 이때, 추출조건은 초록(Abstract) 필드가 널(NULL)이 아닌 것만을 가져온다(단계 1).
- [0043] 이어서, 각 초록의 내용을 문장 단위로 분할하여 큐레이션(curation)을 통해 문장 테이블(Sentence table)로 만들고, 예를 들면, [pubmed id, sentence id, sentence]의 순서로 저장한다(단계 2).
- [0044] 다음으로, 상기 단계에서 얻어진 문장 테이블에 대하여 역인덱스를 구축한다(단계 3).
- [0045] 계속해서, 유전자, 질병, 화합물 동의어 사전을 문장 역인덱스와 비교하여 각각의 차원 인덱스를 구축한다(단계 4).
- [0046] 계속해서, 도 2를 참조하여, 상기한 바와 같이 하여 구축된 인덱스를 저장하는 방법에 대하여 설명한다.
- [0047] 도 2를 참조하면, 도 2는 질병-유전자-화합물의 관계를 추출하기 위한 인덱스 저장구조를 나타내고 있다.
- [0048] 즉, 본 발명의 핵심적인 특징은, 도 2에 나타난 바와 같이, 문장(Sentence)을 각각의 질병, 유전자, 화합물 인덱스의 관점(차원)에서 볼 수 있도록 하는 저장구조에 있는 것이며, 여기서, 이러한 저장구조는, 전문용어로 '스타 스키마(Star Schema)'라고 한다.
- [0049] 더 상세하게는, 도 2에 나타난 바와 같이, 예를 들면, 질병 인덱스에 대하여는, pubmed 아이디, 문장번호, 질병 아이디 및 질병이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 표준 질병명 및 질병에 대한 동의어 정보가 관련지어져 저장된다.
- [0050] 또한, 유전자 인덱스에 대하여도, 마찬가지로, pubmed 아이디, 문장번호, 유전자 아이디 및 유전자 이름, 시작 위치, 종료위치에 대한 정보가 저장되고, 여기에 표준 유전자 및 유전자에 대한 동의어 정보가 관련지어져 저장된다.
- [0051] 아울러, 화합물 인덱스에 대하여는, pubmed 아이디, 문장번호, 합성물 아이디 및 화합물 아이디, 화합물 이름, 시작위치, 종료위치에 대한 정보가 저장되고, 여기에 합성물명 및 합성물에 대한 동의어 정보가 관련지어져 저

장된다.

- [0052] 또한, 그 외에 다른 분석차원에 대하여도, 상기한 바와 같은 내용을 참조하여 필요에 따라 적절히 인덱스 정보만을 추가하면 되므로, 용이하게 또 다른 분석차원을 추가하여 다차원 분석모델을 수립할 수 있다.
- [0053] 이때, 처리 가능한 질의 유형의 예로는, 예를 들면, 검색 대상에 대하여, 한 문장 또는 초록 내에서 원하는 키워드가 포함된 결과를 검색하거나, 또는, 유형별로, 한 가지 또는 두 가지 이상의 조건을 만족하는 문장 또는 초록을 검색하도록 할 수 있다.
- [0054] 다시 말하면, 본 발명은, 문장 전체로서 약 1억여 건에 달하는 문장에 직접 접근하지 않고, 인덱스를 활용하여 검색의 성능과 정확도를 획기적으로 높이는 방법으로서, 문장을 1차원(유전자, 질병, 화합물), 2차원(유전자-질병, 질병-유전자, 유전자-화합물, 화합물-유전자, 질병-화합물, 화합물-질병 관계) 및 3차원(유전자-질병-화합물 관계)으로 분석할 수 있도록 구성한다.
- [0055] 뿐만 아니라, 이러한 저장구조는, 또 다른 분석차원을 용이하게 추가할 수 있는 매우 유연한 구조를 제공할 수 있다.
- [0056] 다음으로, 도 3은 상기한 바와 같은 다차원 분석구조를 적용하여 추출한 기본 검색결과를 나타내는 화면으로, 문장 기반의 키워드 중심의 검색결과를 나타내는 화면이다.
- [0057] 즉, 도 3에 나타낸 바와 같이, 사용자가 검색어를 입력하면, 검색결과를 나타내는 화면에 각 유전자, 질병, 화합물에 대한 색을 각각 적용하여, 시각적인 효과를 줄 뿐만 아니라 사용자가 직관적으로 내용을 이해할 수 있도록 구성한다.
- [0058] 여기서, 각 문장에서의 유전자, 질병, 화합물 정보는 인덱스로부터 가져오도록 한다.
- [0059] 또한, 도 4는 추출된 문장의 전체 초록내용을 제공하는 화면을 나타내고 있다.
- [0060] 즉, 도 4에 나타낸 바와 같이, 사용자가 검색어를 입력하면, 문장을 기준으로 내용을 파악한 다음 전체 초록을 볼 수 있도록 함으로써, 문장을 중심으로 초록 내용을 확인할 수 있도록 구성한다.
- [0061] 계속해서, 도 5는 다차원 분석을 위한 입력화면과 결과화면을 나타내고 있다.
- [0062] 즉, 도 5에 나타낸 바와 같이, 사용자가 화합물에 대한 검색어를 입력하면, 동의어 처리를 통해 관련된 질병에 대한 검색어가 표시되고, 사용자가 그 중 하나를 선택하면 해당하는 화합물과 질병에 대한 검색결과 및 초록이 표시되어, 사용자가 화합물-질병간 관계 분석을 용이하게 수행할 수 있다.
- [0063] 도 6은 도 5에 나타낸 바와 같은 화합물-질병간의 관계를 추출하기 위한 SQL 구조를 나타내고 있다.
- [0064] 즉, 유전자-질병-화합물간의 관계를 추출하기 위한 별도의 프로그램을 작성할 필요 없이, 도 6에 나타낸 바와 같은 SQL을 이용하여 인덱스에 접근하여 필요한 분석을 즉시 수행할 수 있도록 구성한다.
- [0065] 다시 말하면, 상기한 바와 같이 구성된 본 발명의 특징은, 첫째로, 도 3 및 도 4에 나타낸 바와 같이, 문장을 대상으로 키워드 기반 논리검색을 지원하며, 이 검색 기반을 다차원 분석의 최종결과 확인화면으로 활용하는 것과, 둘째로, 도 2에 나타낸 바와 같은 다차원 인덱스 구조를 통해 유전자-질병-화합물 각 관점별로 애드혹(Ad-Hoc) 질의를 수행할 수 있다는 것이다.
- [0066] 계속해서, 도 7 내지 도 10은 상기한 바와 같은 본 발명의 방법을 이용한 다차원 분석 결과에 대한 실제 적용예를 나타내고 있다.
- [0067] 즉, 도 7 내지 도 10에 나타낸 바와 같이, 본 발명에 따르면, 유전자-화합물 관계 분석, 유전자-유전자 관계 분석, 질병-유전자-초록 분석, 질병-유전자 분석과 같이 다양한 다차원 분석이 가능하게 된다.
- [0068] 상기한 바와 같이, 본 발명에 따르면, 고성능의 문장단위의 논리 검색을 지원함으로써, 현재 생명공학 문헌 검색에서 문장단위의 검색을 지원하지 않고 있는 문제를 해결할 수 있다.
- [0069] 또한, 본 발명에 따르면, 검색 결과에서 유전자-질병-화합물 키워드에 대한 하이라이트 기능을 통해 사용자의 직관적 이해도를 향상시킬 있으며, 아울러, 유전자-질병-화합물의 다차원 인덱스를 활용한 유연하고 고성능의 분석 서비스를 제공할 수 있다.
- [0070] 즉, 본 발명은, 예를 들면, 특정 유전자와 관련된 질병리스트 출력, 특정 질병과 관련된 유전자리스트 출력, 특정질병과 유전자가 포함된 초록 가져오기, 특정유전자와 함께 존재하는 유전자 리스트 출력, 특정 유전자와 관

련된 화합물 리스트 출력, 특정 화합물과 관련된 질병 리스트 추력, 특정 신체부위와 관련된 질병리스트 출력, 특정 신체부위와 관련된 화합물 리스트 출력, 특정 종과 관련된 화합물 리스트 출력 등 다양한 경우에 대한 바이오 텍스트 마이닝 서비스를 제공할 수 있다.

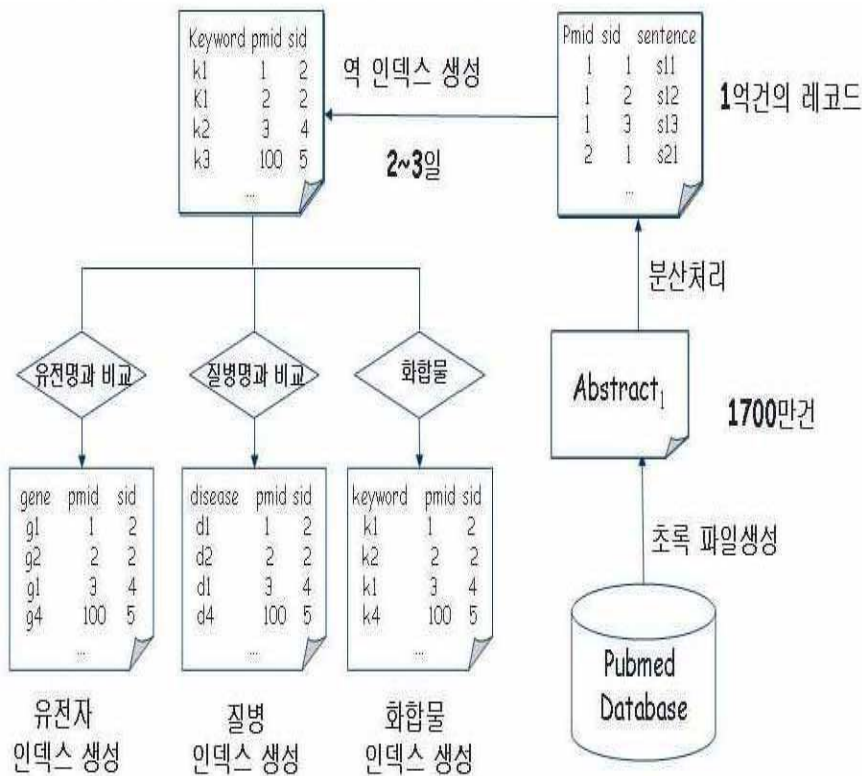
[0071]

이상 상기한 바와 같은 본 발명의 실시예를 통하여 본 발명에 따른 대용량 생명공학 문헌으로부터 유전자-질병-화합물 관계를 추출하는 방법의 상세한 내용에 대하여 설명하였으나, 본 발명은 상기한 실시예에 기재된 내용으로만 한정되는 것은 아니며, 따라서 본 발명은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 설계상의 필요 및 기타 다양한 요인에 따라 여러 가지 수정, 변경, 결합 및 대체 등이 가능한 것임은 당연한 일이라 하겠다.

도면

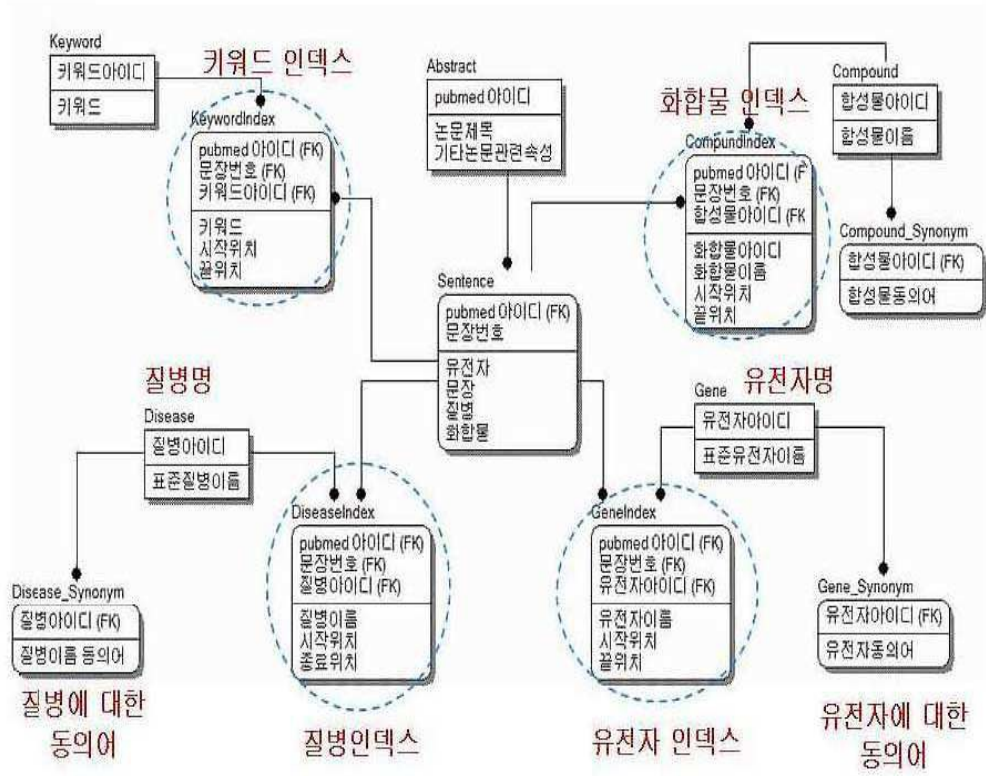
도면1

대용량 문헌으로부터 질병-유전자-화합물에 대한 다차원 인덱스 구축절차



도면2

질병-유전자-화합물의 관계를 추출하기 위한 저장구조



도면3

문장기반의 키워드 중심의 검색 결과

In Sentence BRCA1 AND "breast cancer" AND (bind* OR interact*) Search DB search

Found 81 documents(s) (in 265 milliseconds) that matched query 'BRCA1 AND "breast cancer" AND (bind* OR interact*)':

- 18543242 | 7 | An interaction between BRCA1 and HIF-1alpha was found in human Breast cancer Cells
- 18071589 | 0 | Cofactor of BRCA1 (COBRA1) was first identified as a protein that binds to the Breast cancer susceptibility gene product BRCA1.
- 15205325 | 3 | BRCA1 disease-associated mutations abrogated the interaction in yeast, and BRCA1 failed to interact with MEKK3 in BRCA1 mutant HCC1937 Breast cancer cells.
- 13679996 | 3 | Here we discuss the role of BRCA1 and NER in Breast cancer and the interactions of BRCA1 with p53 in Breast tumorigenesis and suggest approaches for risk assess cancer.
- 15199145 | 8 | Our findings support a novel tumor suppressor function of BRCA1 that is relevant to Breast cancer and identify a potential interactive risk factor for women with BRCA1
- 18501503 | 2 | Here we show that BRCA1 interacts with p53 in vivo in Breast cancer cells, through another BRCA1 domain (amino Acids 772-1292).
- 8944023 | 3 | The BARD1/BRCA1 interaction is disrupted by BRCA1 missense mutations that segregate with Breast cancer susceptibility, indicating that BARD1 may be involved in medi
- 11733219 | 2 | This diverse group of disorders is thought to be linked through protein interactions with the Breast cancer tumor susceptibility gene product, BRCA1.
- 18519686 | 1 | In this study, we show that the BRCT mutant of BRCA1, R1699W, which is associated with increased risk of Breast cancer, is unable to interact with RAP80.
- 11916966 | 1 | We observed a novel endogenous association of BRCA1 with NMI (N-MYC-interacting protein) in Breast cancer cells.
- 17889711 | 1 | Human RAD51, known to function in DNA repair, interacts with a number of Proteins implicated in Breast cancer (BC), including BRCA1 and BRCA2
- 18109739 | 1 | Previously, the Breast cancer susceptibility gene product BRCA1 was found to interact with and inhibit the transcriptional activity of estrogen receptor-alpha.
- 16357529 | 0 | The BACH1 helicase was initially identified by its direct binding to BRCA1 and, thus, was linked to hereditary Breast cancer.
- 11562180 | 4 | Also, additional Breast cancer genes can be identified by further characterization of the function of BRCA1 and BRCA2 and their interacting Proteins.
- 10868324 | 11 | This interaction between ATM and BRCA1 argues in favor of the involvement of particular aspects of ATM function in Breast cancer predisposition.
- 18501503 | 4 | These results suggest that a novel domain of BRCA1 may interact with p53 in Breast cancer cells.

도면4

문장을 중심으로 초록 내용 확인 화면

In Sentence Search DB search

Found 81 documents(s) (in 265 milliseconds) that matched query BRCA1 AND "breast cancer" AND (bind* OR interact*):

An interaction between BRCA1 and HIF-1alpha was found in human Breast cancer Cells

A recent study of Breast cancer patients with and without BRCA1 gene mutations found significantly lower levels of VEGF in serum from patients with BRCA1 mutations (Tarnowski, B., Chudecka-Glaz, A., Gorski, B., and Rzepta-Gorska, I. (2004) Breast Cancer Res.Treat. 88, 287-288). Here, we describe a possible mechanistic explanation for this correlation. Because hypoxia in tumors stimulates VEGF expression and secretion we hypothesized that altered BRCA1 protein levels in Breast tumors could affect hypoxia-stimulated VEGF promoter activity. This possibility was tested in cells transfected with various combinations of expression plasmids for BRCA1, BRCA1 specific inhibitory RNAs (BRCA1-siRNAs), HIF-1alpha, and a VEGF promoter-reporter and then incubated in normoxia (21% O2) or hypoxia (0.1% O2). As predicted, increased BRCA1 levels enhanced both hypoxia-stimulated VEGF promoter activity and the amounts of VEGF mRNA, as determined by semiquantitative RT-PCR and quantitative real time PCR. Using the CHIP assay, we discovered that BRCA1 could be recruited to the endogenous human VEGF promoter along with HIF-1alpha following hypoxia. **An interaction between BRCA1 and HIF-1alpha was found in human Breast cancer Cells.** We also found that hypoxia-stimulated VEGF promoter activity and secretion was reduced in cells containing reduced amounts of endogenous BRCA1 protein (obtained by transfecting with BRCA1 siRNAs). A mechanistic explanation for these effects is provided by our finding a reduced half-life and reduced accumulation of HIF-1alpha in hypoxic cells transfected with BRCA1-siRNAs and that proteasome inhibitors blocked these effects of BRCA1-siRNAs. Thus, our results suggest that normal amounts of BRCA1 function in hypoxia to regulate HIF-1alpha stability, probably by interacting with HIF-1alpha.

도면5

화합물-질병간의 관계 분석 위한 입력화면과 결과화면

화합물:

Search

동의어 처리

term	COUNT (**)
Asthma	126
Stroke	77
Thrombosis	42
Asthma, Aspirin-Inducer	42
Urticaria	38
Infection	35
Pain	34
Myocardial Infarction	30
Hypersensitivity	22
Ulcer	19
Syndrome	19
Polyps	13
Death	13
Coronary Disease	12
Recurrence	11
Arthritis	11
Angioedema	10
Atherosclerosis	9
Coronary Artery Disease	9
Inflammation	8
Arthritis, Rheumatoid	8
Nasal Polyps	8
Flushing	7

3165682 | 11 | Aspirin was not effective in preventing Thrombosis

17765703 | 6 | Aspirin was given as a prophylaxis for Thrombosis

9054765 | 13 | Stent Thrombosis was resistant to Heparin and Aspirin

9106227 | 10 | Trials of low dose Aspirin to prevent Thrombosis are encouraged

14765622 | 8 | Aspirin is generally added in case of arterial Thrombosis

15804471 | 2 | While off Aspirin, these patients may be at risk of Thrombosis

10657017 | 1 | Yet, Aspirin is unable to prevent Thrombosis in all patients

7051035 | 6 | Aspirin reduced the frequency of arteriovenous shunt Thrombosis

12176945 | 2 | Late Thrombosis was attenuated with 6 months of Aspirin and clopidogrel

11152927 | 0 | To prevent arterial Thrombosis, abciximab is administered together with Aspirin

3431418 | 3 | Aspirin helps prevent Thrombosis, but not as well as Heparin

1072621 | 0 | Aspirin and Ticlopidine are used to prevent arterial Thrombosis

9110 | 클릭 Aspirin and Dipyridamole DO not decrease the frequency of shunt Thrombosis

1650146 | 1 | Thrombosis is effectively prevented by antiplatelet therapy with Aspirin and clopidogrel

초록전체 보여주기

3165682 | 11 | Aspirin was not effective in preventing Thrombosis

17765703 | 6 | Aspirin was given as a prophylaxis for Thrombosis

9054765 | 13 | Stent Thrombosis was resistant to Heparin and Aspirin

9106227 | 10 | Trials of low dose Aspirin to prevent Thrombosis are encouraged

14765622 | 8 | Aspirin is generally added in case of arterial Thrombosis

15804471 | 2 | While off Aspirin, these patients may be at risk of Thrombosis

10657017 | 1 | Yet, Aspirin is unable to prevent Thrombosis in all patients

7051035 | 6 | Aspirin reduced the frequency of arteriovenous shunt Thrombosis

12176945 | 2 | Late Thrombosis was attenuated with 6 months of Aspirin and clopidogrel

11152927 | 0 | To prevent arterial Thrombosis, abciximab is administered together with Aspirin

3431418 | 3 | Aspirin helps prevent Thrombosis, but not as well as Heparin

A variety of anticoagulation and antiplatelet aggregation agents are in use clinically, and in applications to microvascular surgery, a clear choice for the most effective therapy has not been determined. Models of vascular trauma combined with microvascular anastomosis have been developed for comparing the efficacy of these agents in maintaining vascular patency in a rat model of microvascular Thrombosis. Heparin effectively prevents occlusion in both Arteries and Veins, at clinical levels of administration. Aspirin helps prevent Thrombosis, but not as well as Heparin. These results support the beneficial effect of antithrombotic drug therapies, and suggest a more potent role of the Heparin-inhibited Fibrin clot over platelet aggregation in creating thrombotic occlusion of small vessels.

도면6

화합물-질병간의 관계를 추출하기 위한 SQL

```

SELECT MT.term, COUNT(*)
FROM (SELECT Cl.pmid, Cl.sid, term
      FROM mesh_term MT, ChemicalIndex CI
      WHERE MT.no = Cl.cid AND
           MT.term = 'Aspirin' LIMIT 5000) AA,
      DiseaseIndex DI, mesh_term MT
WHERE AA.pmid = DI.pmid AND AA.sid = DI.sid AND DI._mid=MT.no
GROUP BY MT.term
ORDER BY COUNT(*) DESC;
    
```

도면7

유전자-화합물 관계 분석 화면

유전자:

Search

동의어 처리

초록전체 보여주기

2443279 | 2 | Ceruloplasmin (CP) and alpha2-macroglobulin (AZM) were measured by ELISAs thereby avoiding concentration of BALF.

1722474 | 4 | Age did not contribute to the AZM levels but did have a negative effect on the HP, Ceruloplasmin (CP) and fibrinogen (Fib) levels.

3102187 | 2 | Nevertheless a highly significant rise of alpha 2 macroglobulin (AZM) and Ceruloplasmin (CPL) levels in this subgroup was stated.

17620124 | 20 | Following Proteins were analyzed:CRP, AGR, alpha1-antitrypsin (ACT), Transferrin, Ceruloplasmin (CP), alpha1-antitrypsin, haptoglobin and alpha2-macroglobulin (AZM).

15327016 | 15 | Slight age-related differences were observed, irrespective of the region, in IGA and Ceruloplasmin (increase) and AZM (decrease).

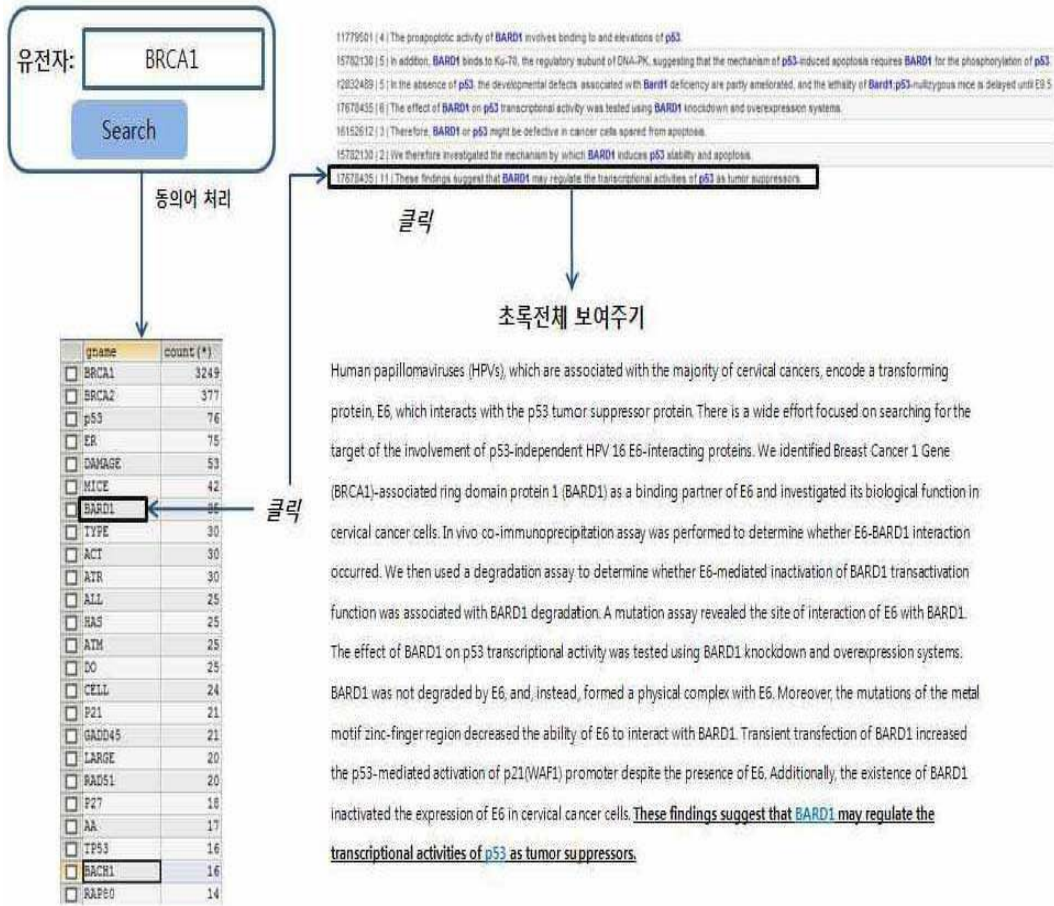
클릭

The I.U.C. Committee on **reference proteins** has been investigating regional differences for commonly assayed **reference proteins** to determine whether universal reference intervals can be applied. As a part of this study, we launched an Asian project analyzing the concentrations of 13 serum proteins whose values are standardized to CRM470, and five newer analytes: retinol-binding protein (RBP), Cystatin C (CysC), light-chain-kappa (L-kappa), and light-chain-lambda (L-lambda) in Tokyo, Seoul, Kuala Lumpur, Hong Kong, Taipei and Shanghai; serum samples were collected from 146 to 415 apparently healthy individuals with nearly equal gender ratios. All assays were performed in Tokyo on a Behring Nephelometer II (BN II). Seven chemical analytes (aspartate aminotransferase (AST), Alanine aminotransferase (ALT), gamma-glutamyltransferase (gammaGT), Creatinine, total Cholesterol (TC), Triglycerides (TG) and high-density lipoprotein Cholesterol (HDL-C)) were also measured. These results were used for excluding individuals with possible latent clinical disorders. Positive acute phase reactants were consistently lower, and negative ones were higher, in Tokyo than those in other cities. The most conspicuous difference was observed in C-Reactive Protein (CRP). There were no regional differences in Transferrin, albumin, or CysC. Creatinine was much lower in Tokyo despite comparable CysC levels. ALT and gammaGT were higher in Shanghai, Taipei and Seoul; gammaGT and TG were higher in Shanghai and HDL-C was higher in Tokyo. Gender-related differences in reference intervals were observed for immunoglobulin (IgM), haptoglobin, RBP, Transferrin, alpha2-macroglobulin (AZM), transferrin, alpha1-acid glycoprotein, CysC, and C4 in all cities. Slight age-related differences were observed, irrespective of the region, in IGA and Ceruloplasmin (increase) and AZM (decrease). Environmental factors and lifestyle seem to have a GREAT influence on many commonly measured analytes.

Immunoglobulin A	15
Transferrin	15
Ceruloplasmin	15
Amyloid	11
Lipoprotein(a)	10
Antibodies	10
Proteins	10
Immunoglobulins	9
Protein C	9
Fibrinogen	9
Orosomucoid	7
C-Reactive Protein	7
Cytokines	5
Prothrombin	5
Zinc	5
Turpentine	4
Serum Albumin	4
alpha 1-Antitrypsin	4
Antithrombin III	4
Antigens	4

도면8

유전자-유전자 분석 화면



도면9

질병-유전자-초록 분석화면

질병: Ovarian Neoplasms

유전자: p53

Search

동의어 처리

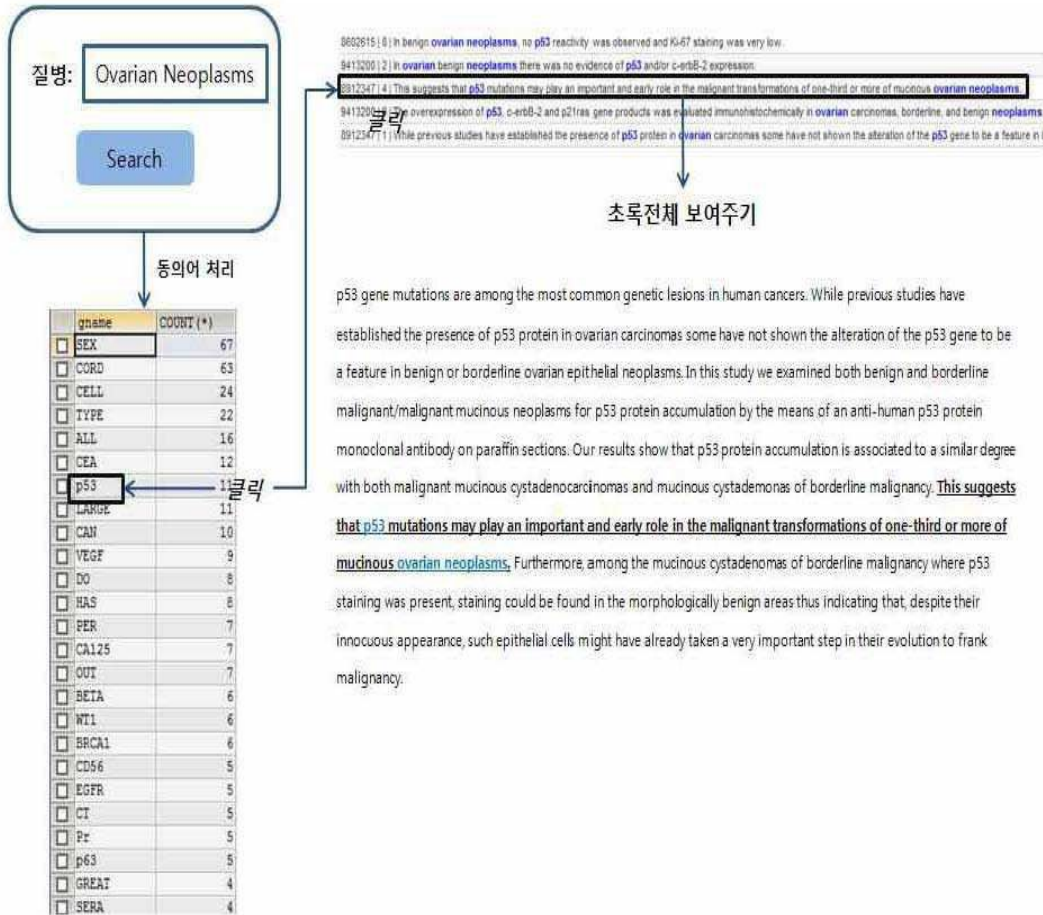
pmid
<input type="checkbox"/> 8640675
<input type="checkbox"/> 11894392
<input type="checkbox"/> 9556793
<input type="checkbox"/> 8602615
<input type="checkbox"/> 10320143
<input type="checkbox"/> 9987671
<input type="checkbox"/> 9413200
<input type="checkbox"/> 9166285
<input type="checkbox"/> 15985106
<input type="checkbox"/> 16776856
<input type="checkbox"/> 15806512
<input type="checkbox"/> 10374771
<input type="checkbox"/> 8912347
<input type="checkbox"/> 9010031

클릭

The phenotypic variability of epithelial ovarian neoplasms correlates with a diversity of changes on the molecular level. Invasive serous and undifferentiated ovarian carcinomas are characterized by p53 mutations, extensive loss of genetic material of chromosome 17 and complex changes on many other chromosomes. These alterations are seen only in a minority of mucinous and endometrioid carcinomas, mainly in advanced stages. Overexpression of bcl-2 is seen most frequently in endometrioid carcinomas (ca. 90% of cases), which in addition show microsatellite instability in around a third of cases, as has been described in endometrioid endometrial carcinomas. KRAS mutations are characteristic for mucinous LMP tumors (borderline tumors) and mucinous carcinomas (40-50% of cases). Furthermore, KRAS mutations have been described in around a third of serous LMP tumors, which in addition show microsatellite instability in up to 40% of cases. Serous LMP tumors never harbour complex chromosomal aberrations.

도면10

질병-유전자 분석 화면



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 8의 2~3행

【변경전】

상기 질병 인덱스, 상기 유전자인덱스 및 상기 화합물 인덱스

【변경후】

질병 인덱스, 유전자 인덱스 및 화합물 인덱스