



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0079641
(43) 공개일자 2015년07월08일

- (51) 국제특허분류(Int. Cl.)
G06F 19/24 (2011.01) G01N 33/50 (2006.01)
G01N 33/68 (2006.01) G06F 19/18 (2011.01)
- (52) CPC특허분류
G06F 19/24 (2013.01)
C12Q 1/68 (2013.01)
- (21) 출원번호 10-2015-7010816
- (22) 출원일자(국제) 2013년02월12일
심사청구일자 없음
- (85) 번역문제출일자 2015년04월24일
- (86) 국제출원번호 PCT/JP2013/053188
- (87) 국제공개번호 WO 2014/050160
국제공개일자 2014년04월03일
- (30) 우선권주장
JP-P-2012-211921 2012년09월26일 일본(JP)

- (71) 출원인
고쿠리츠켄큐카이하츠호진 카가쿠기쥬츠신코키코
일본 사이따마켄 가와구쨌시 혼쇼 4쵸메 1방 8고
- (72) 발명자
아이하라 가즈유키
일본 도쿄도 1538505 메구로쿠 고마바 4-6-1
첸 루오난
일본 효고켄 6580032 고베시 히가시나다쿠 고요쇼
나카 7-2-1-5-808
(뒷면에 계속)
- (74) 대리인
김태홍, 김진희

전체 청구항 수 : 총 11 항

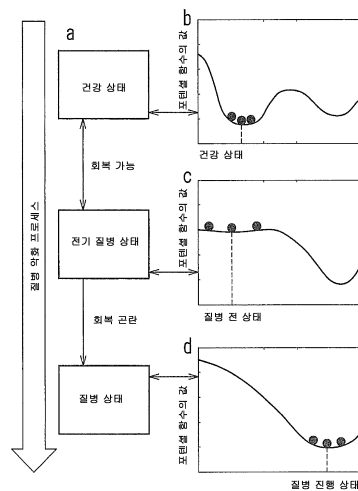
(54) 발명의 명칭 동적 네트워크 바이오마커의 검출 장치, 검출 방법 및 검출 프로그램

(57) 요약

건강 상태에서부터 질병 상태로의 천이를 나타내는 질병 전 상태의 진단에 이용하는 것이 가능한 바이오마커의 후보를 검출하는 검출 장치, 검출 방법 및 검출 프로그램을 제공한다.

측정 대상으로부터 복수의 생체 샘플을 상이한 시간점에서 채취하고, 채취한 생체 샘플을 측정하여 얻어진 측정 데이터를 집계하여 통계 데이터를 취득한다. 그리고, 하이 스트루프ット 데이터의 취득 처리(s1)와, 디퍼렌셜 생체 분자의 선출 처리(s2)와, 클러스터화 처리(s3)와, DNB의 후보의 선출 처리(s4)와, 유의성 분석에 의한 DNB의 판정 처리(s5)를 행한다.

대표도 - 도1



(52) CPC특허분류

G01N 33/50 (2013.01)

G01N 33/68 (2013.01)

G06F 19/18 (2013.01)

(72) 발명자

리우 루이

중국 광둥주 510640 광저우 톈허 디스트릭트 381
우산 로드 사우스 차이나 유니버시티 오브 테크놀
로지 룸 4225 4번가 빌딩

리우 지펑

중국 산둥 250100 지난 홍지아로우 커뮤니티 슈푸
로드 넘버 1 산둥 유니버시티 패컬티 레지덴셜 아
레아 1-2-501

리 메이이

중국 상하이 200031 웨양 로드 넘버 319 종케 빌딩

명세서

청구범위

청구항 1

생체에 관한 측정치에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 검출 장치로서,

상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 복수의 인자 항목을 복수의 클러스터로 분류하는 분류 수단과,

분류한 각 클러스터로부터, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화 및 각 인자 항목간의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 미리 설정된 선출 조건에 해당하는 클러스터를 선출하는 선출 수단과,

선출한 클러스터에 포함되는 인자 항목을 바이오마커의 후보로서 검출하는 검출 수단

을 포함하는 것을 특징으로 하는 검출 장치.

청구항 2

제1항에 있어서,

상기 선출 수단은,

클러스터 내의 각 인자 항목간의 측정 데이터의 상관을 나타내는 값의 평균값을 제1 지수로서 산출하는 수단과,

클러스터 내의 인자 항목의 측정 데이터와 상기 클러스터 외의 인자 항목의 측정 데이터 간의 상관을 나타내는 값의 평균값을 제2 지수로서 산출하는 수단과,

클러스터 내의 각 인자 항목에 대해서 측정 데이터의 표준 편차의 평균값을 제3 지수로서 산출하는 수단

을 포함하고,

복수의 클러스터 중에서, 제1 지수, 제2 지수 및 제3 지수에 기초하여, 바이오마커로 해야 하는 인자 항목을 포함하는 클러스터를 선출하는 것을 특징으로 하는 검출 장치.

청구항 3

제2항에 있어서,

상기 선출 수단은, 상기 제1 지수와, 상기 제2 지수와, 상기 제3 지수의 역수와의 곱에 기초한 종합 지수가 최대인 클러스터를 선택하는 것을 특징으로 하는 검출 장치.

청구항 4

제2항 또는 제3항에 있어서,

상기 각 인자 항목의 각각의 측정 데이터가, 유의성을 가지고 경시적으로 변화하고 있는지의 여부를 검정하는 디퍼렌셜 검정 수단을 더 포함하고,

상기 분류 수단은, 경시적 변화에 유의성이 있으면 검정된 인자 항목에 대해서 분류하는 것을 특징으로 하는 검출 장치.

청구항 5

제4항에 있어서,

상기 디퍼렌셜 검정 수단은, 각 인자 항목의 측정 데이터, 및 인자 항목 및 시계열마다 미리 설정되어 있는 참조 데이터와의 비교 결과에 기초하여, 유의성에 관한 검정을 행하는 것을 특징으로 하는 검출 장치.

청구항 6

제5항에 있어서,

각 인자 항목에 대해서, 대응하는 참조 데이터의 표준 편차의 평균값을 나타내는 참조 표준 편차, 및 인자 항목 간의 상관을 나타내는 값의 평균값을 나타내는 참조 상관값을 산출하는 수단을 더 포함하고,

상기 검출 수단은, 상기 제1 지수가 상기 참조 표준 편차에 비해서 유의성을 가지고 증대하며, 상기 제2 지수가 상기 참조 상관값에 비해서 유의성을 가지고 감소하고, 또한, 상기 제3 지수가 상기 참조 표준 편차에 비해서 유의성을 가지고 증대한 경우에, 상기 클러스터에 포함되는 항목을 바이오마커의 후보로서 검출하는 것을 특징으로 하는 검출 장치.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서,

상기 검출 수단은,

클러스터에 포함되는 복수의 인자 항목의 유의성을, 측정 데이터의 통계값에 기초하여 검정하는 수단을 포함하고,

유의성이 있는 경우에, 클러스터에 포함되는 항목을 바이오마커의 후보로서 검출하는 것을 특징으로 하는 검출 장치.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서,

상기 복수의 인자 항목 중 어느 하나는, 유전자에 관한 측정 항목, 단백질에 관한 측정 항목, 대사물에 관한 측정 항목, 또는 생체로부터 얻어지는 화상에 관한 측정 항목인 것을 특징으로 하는 검출 장치.

청구항 9

생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 검출 장치를 이용한 검출 방법으로서,

상기 검출 장치는,

상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 복수의 인자 항목을 복수의 클러스터로 분류하는 분류 단계와,

분류한 각 클러스터로부터, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화 및 각 인자 항목간의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 미리 설정된 선출 조건에 해당하는 클러스터를 선출하는 선출 단계와,

선출한 클러스터에 포함되는 인자 항목을 바이오마커의 후보로서 검출하는 검출 단계를 실행하는 것을 특징으로 하는 검출 방법.

청구항 10

생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 검출 방법으로서,

상기한 시간점에서 채취한 복수의 생체 샘플의 각각으로부터 측정된 하이 스루풋 데이터 중에서, 디퍼렌셜 생체 분자를 산출하는 분자 스크리닝 단계와,

상관이 높은 생체 분자끼리를 하나의 클러스터로 하도록, 상기 분자 스크리닝 단계에서 선출한 상기 디퍼렌셜 생체 분자를 복수의 클러스터로 분류하는 클러스터화 단계와,

상기 클러스터화 단계에서 얻어진 복수의 클러스터 중에서, 클러스터 내의 생체 분자 간의 상관의 증대, 클러스터 내의 생체 분자의 표준 편차의 증대, 및 클러스터 내의 생체 분자와 다른 생체 분자 간의 상관의 저감이 가장 현저한 클러스터를, 상기 바이오마커의 후보로서 선취하는 후보 선택 단계와,

상기 후보 선택 단계에서 선출한 바이오마커의 후보가 상기 바이오마커인지의 여부를, 유의성 검정에 의해 판정하는 판정 단계

를 실행하는 것을 특징으로 하는 검출 방법.

청구항 11

컴퓨터에, 생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 처리를 실행시키는 검출 프로그램으로서,

컴퓨터에,

상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 복수의 인자 항목을 복수의 클러스터로 분류하는 분류 단계와,

분류한 각 클러스터로부터, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화 및 각 인자 항목간의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 미리 설정된 선출 조건에 해당하는 클러스터를 선출하는 선출 단계와,

선출한 클러스터에 포함되는 인자 항목을 바이오마커의 후보로서 검출하는 검출 단계

를 실행시키는 것을 특징으로 하는 검출 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은, 생체에 관한 측정에 의해 얻어진 복수의 인자 항목의 측정 데이터에 기초하여, 상기 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 검출 장치, 검출 방법 및 검출 프로그램에 관한 것이다.

배경 기술

[0002] 여러가지 연구 결과에 따르면, 많은 질병, 특히 복잡한 질병의 악화 진행의 프로세스는, 기후 시스템, 생태 시스템, 경제 시스템 등의 시스템과 마찬가지로, 어떤 임계 역치를 넘은 시점, 소위 분기점에 달하면, 상태 천이가 돌연 발생하여, 건강 상태로부터 급격하게 질병 상태로 변화한다(예컨대, 비특허문헌 1~5 참조). 이러한 복잡 질병의 동적 메커니즘에 관한 연구에 있어서, 질병 악화(예컨대, 천식 발작, 암의 발병)의 진행 프로세스를, 시간에 의존한 비선형 동역학 시스템으로서 모델화하고, 모델화한 시스템을 관측함으로써, 분기점에서의 상전이로 병이 급격하게 악화하는 것이 이미 판명되어 있다(비특허문헌 1, 6 참조).

[0003] 도 1은 질병의 진행 프로세스를 개념적으로 나타내는 설명도이다. 도 1의 a는 질병의 진행 프로세스를 모식적으로 나타내고 있다. 도 1의 b, c 및 d는 진행 프로세스의 과정에 있어서, 전술한 모델화한 시스템의 안정성을 포텐셜 함수로서 나타내고, 횡축에 경과를 나타내는 시각을 들며, 종축에 포텐셜 함수의 값을 들어 개념적으로 나타낸 그래프이다. 도 1의 a에 나타내는 바와 같이, 질병 악화의 진행 프로세스는, 정상 상태(건강 상태), 질병 전 상태, 질병 상태로서 나타낼 수 있다. 정상 상태에 있어서, 시스템은 안정되어, 도 1의 b에 검은 동그라미의 위치로서 나타내는 바와 같이, 포텐셜 함수의 값이 최소값으로 되어 있다. 질병 전 상태에 있어서, 시스템은, 도 1의 c에 검은 동그라미의 위치로서 나타내는 바와 같이, 포텐셜 함수의 값이 높게 되어 있다. 따라서, 외란의 영향을 받기 쉬운 상태이며, 작은 외란을 받는 것만으로 상전이하여 버리는 분기점의 부근, 즉, 정상 상태의 한계에 위치하고 있다. 단, 상기 질병 전 상태는, 적절한 처치에 의해, 용이하게 정상 상태로 회복될 수 있다. 한편, 질병 상태에 있어서, 시스템은 안정되어, 도 1의 d에 검은 동그라미의 위치로서 나타내는 바와 같이, 포텐셜 함수의 값이 대국적 최소값으로 되어 있다. 그 때문에, 정상 상태에서부터 상전리로 생긴 이 질병 상태가, 정상 상태로 회복되는 것은 곤란하다.

[0004] 따라서, 혹시 질병 전 상태를 검지하여, 질병 상태로 천이하기 전에, 질병 상태로 천이하고 있는 것을 환자에게 예고할 수 있으면, 적절한 조치를 취함으로써, 환자를 질병 전 상태에서부터 정상 상태로 회복시킬 수 있는 가능성이 높다.

[0005] 즉, 분기점(임계 역치)을 검출할 수 있으면, 임계 천이의 예측이 가능하여, 병의 조기 진단을 실현할 수 있다.

[0006] 또한, 종래부터 질병 상태를 진단하는 방법으로서, 바이오마커가 이용되고 있다. 종래부터 이용되고 있는 통상

의 바이오마커는, 생체로부터 채취된 혈청, 뇨 등의 체액, 또한, 조직에 포함되는 분자 레벨의 DNA, RNA, 단백질, 대사물 등이며, 생체 내의 생물학적 변화를 정량적으로 파악할 수 있는 지표이다. 종래의 바이오마커에 의한 질병의 진단 방법은, 노멀 샘플(건강 상태에서 채취한 샘플)로부터 추출한 바이오마커와 이상 샘플(질병 상태에서 채취한 샘플)로부터 추출한 바이오마커를 비교함으로써, 병을 진단하는 방법이다.

선행기술문헌

비특허문헌

[0007]

(비특허문헌 0001) 비특허문헌 1: 베네가스·제이·지(Venegas, J.G.) 등 편저, 「과피 변화에 대한 전주곡과 같은 천식의 자기 조직(Self-organized patchiness in asthma as a prelude to catastrophic shifts.), (영국), 네이처(Nature), 네이처 퍼블리싱 그룹(Nature Publishing Group), 2005년, 434권, p.777-782

(비특허문헌 0002) 비특허문헌 2: 맥셰리·피·이(McSharry, P.E.), 스미스·엘·에이(Smith, L.A.), 타라센코·엘(Tarassenko, L) 편저, 「뻗성 발작의 예측: 비선형 방법이 적절인가.(Prediction of epileptic seizures: are nonlinear methods relevant.)」, (영국), 네이처·메디신(Nature Medicine), 네이처 퍼블리싱 그룹(Nature Publishing Group), 2003년, 9권, p.241-242

(비특허문헌 0003) 비특허문헌 3: 로베르토·피·비(Roberto, P.B.), 엘리세오·지(Eliseo, G.), 조세프·에프·씨(Jose f, C.), 「물류 회귀의 변화점 추정을 위한 전이 모델(Transition models for change-point estimation in logistic regression.)」, (미국), 의학에 있어서의 통계(Statistics in Medicine), 와일리-블랙웰(Wiley-Blackwell), 2003년, 22권, p.1141-1162

(비특허문헌 0004) 비특허문헌 4: 궁·에스(Paek, S.) 등 편저, 「음향 신경종에 대한 감마 나이프 수술 후, 청각 보존(Hearing preservation after gamma knife stereo tactic radiosurgery of vestibular schwannoma.)」, (미국), 캔서(Cancer), 와일리-블랙웰(Wiley-Blackwell), 2005년, 1040권, p.580-590

(비특허문헌 0005) 비특허문헌 5: 리우·제이·케이(Liu, J.K.), 로빗·알·엘(Rovit, R.L.), 쿼드웰·더블유·티(Couldwell, W.T.), 「뇌하수체 졸중(Pituitary Apoplexy)」, (미국), 세미나즈·인·뉴로서저리(Seminars in neurosurgery), 티메(Thieme) 2001년, 12권, p.315-320

(비특허문헌 0006) 비특허문헌 6: 다나카·지(Tanaka, G.), 츠모토·케이(Tsumoto, K.), 츠지·에스(Tsuji, S.), 아이하라·케이(Aihara, K.) 편저, 「전립선암에 대한 간헐적인 호르몬 요법의 하이브리드 시스템 모델의 분기점 분석.(Bifurcation analysis on a hybrid systems model of intermittent hormonal therapy for prostate cancer.)」, (미국), 피지컬·리뷰(Physical Review), 미국 물리학회, 2008년, 237권, p.2616-2627

발명의 내용

해결하려는 과제

[0008]

그러나, 복잡 질병의 경우, 임계 천이의 예측은 매우 곤란하다. 그 이유는 이하와 같다.

[0009]

첫 번째로, 질병 전 상태는, 정상 상태의 경계이며, 분기점에 달하기 전에, 현저한 변화가 생기기 어렵다. 그 때문에, 종래의 바이오마커, 스냅 샷 측정 등의 방법에 따른 진단에서는, 정상 상태와 질병 전 상태를 구별하는 것이 어렵다.

[0010]

두 번째로, 여러가지 연구가 이루어져 있는 중에, 분기점을 예측하기 위한 조기 진단용의 경고 신호를 정밀도 높게 검출할 수 있는 신뢰성이 높은 질병 모델이 아직 개발되어 있지 않다. 특히, 동일한 질병이라도, 개인에 따라, 질병 악화의 진행 프로세스가 상이하기 때문에, 모델 베이스의 진단 방법은, 성공할 확률이 낮다.

[0011]

세 번째로, 질병 전 상태의 검출 대상은 환자이며, 통상, 1인의 환자로부터 얻어지는 샘플의 수가 한정되어 있기 때문에, 장기간에 걸쳐, 예측에 필요한 충분한 샘플을 채취하는 것이 곤란하다.

[0012]

더구나, 종래의 바이오마커에 의한 질병의 진단 방법은, 정상 상태와 질병 상태를 비교함으로써 진단을 행하는 방법이며, 진단된 시점에서는, 환자는 이미 질병 상태에 빠져, 이전의 정상 상태로 되돌아가는 것은 어렵다.

[0013]

그에 대하여, 본 발명은, 질병 상태로 천이하기 전의 질병 전 상태를 검출할 수 있는 검출 장치, 검출 방법 및

검출 프로그램을 제공하며, 또한, 질병 모델이 불필요하고, 적은 수의 생체 샘플만으로 진단을 보조하는 것이 가능한 검출 장치 등을 제공하는 것을 목적으로 하고 있다.

과제의 해결 수단

- [0014] 상기 목적을 달성하기 위해, 본 발명에 따른 검출 장치는, 생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 검출 장치로서, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 복수의 인자 항목을 복수의 클러스터로 분류하는 분류 수단과, 분류한 각 클러스터로부터, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화 및 각 인자 항목간의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 미리 설정된 선출 조건에 해당하는 클러스터를 선출하는 선출 수단과, 선출한 클러스터에 포함되는 인자 항목을 바이오마커의 후보로서 검출하는 검출 수단을 구비하는 것을 특징으로 한다.
- [0015] 상기 특징을 갖는 검출 장치는, 정상 상태에서부터 질병 상태로 천이하기 전의 질병 전 상태를 나타내는 경고 신호가 되는 바이오마커의 후보를 검출할 수 있다. 또한, 바이오마커를 특정할 수 있으면, 검출 대상으로부터 적은 수의 샘플을 채취하는 것만으로 질병 전 상태를 판별할 수 있다.
- [0016] 또한, 본 발명에 따른 검출 장치에 있어서, 상기 선출 수단은, 클러스터 내의 각 인자 항목간의 측정 데이터의 상관관계를 나타내는 값의 평균값을 제1 지수로서 산출하는 수단과, 클러스터 내의 인자 항목의 측정 데이터와 상기 클러스터 외의 인자 항목의 측정 데이터 간의 상관관계를 나타내는 값의 평균값을 제2 지수로서 산출하는 수단과, 클러스터 내의 각 인자 항목에 대해서 측정 데이터의 표준 편차의 평균값을 제3 지수로서 산출하는 수단을 포함하고, 복수의 클러스터 중에서, 제1 지수, 제2 지수 및 제3 지수에 기초하여, 바이오마커로 해야 하는 인자 항목을 포함하는 클러스터를 선출하는 것을 특징으로 한다.
- [0017] 따라서, 상기 제1 지수, 제2 지수, 제3 지수를 이용함으로써, 각 클러스터의 특성을 정량적 평가할 수 있어, 바이오마커를 용이하게 선출할 수 있다.
- [0018] 또한, 본 발명에 따른 검출 장치에 있어서, 상기 선출 수단은, 상기 제1 지수와, 상기 제2 지수와, 상기 제3 지수의 역수의 곱에 기초한 종합 지수가 최대인 클러스터를 선택하는 것을 특징으로 한다.
- [0019] 따라서, 종합 지수에 기초하여 클러스터를 선택함으로써, 바이오마커의 후보가 되는 인자 항목의 신뢰성을 높일 수 있다.
- [0020] 또한, 본 발명에 따른 검출 장치는, 상기 각 인자 항목의 각각의 측정 데이터가, 유의성을 가지고 경시적으로 변화하고 있는지의 여부를 검정하는 디퍼렌셜 검정 수단을 더 구비하고, 상기 분류 수단은, 경시적 변화에 유의성이 있다고 검정된 인자 항목에 대해서 분류하는 것을 특징으로 한다.
- [0021] 따라서, 경시적으로 현저한 변화를 나타낸 인자 항목을 선출함으로써, 방대한 측정 데이터로부터, 바이오마커의 후보를 효율적으로 검출할 수 있다.
- [0022] 또한, 본 발명에 따른 검출 장치는, 상기 디퍼렌셜 검정 수단은, 각 인자 항목의 측정 데이터, 및 인자 항목 및 시계열마다 미리 설정되어 있는 참조 데이터와의 비교 결과에 기초하여, 유의성에 관한 검정을 행하는 것을 특징으로 한다.
- [0023] 따라서, 검출 대상이 되는 상기 복수의 인자 항목에 관한 측정 데이터 외에, 참조 기준이 되는 생체 샘플을 참조 데이터로서 취득함으로써, 측정 데이터와 참조 데이터를 비교하여, 외란을 배제하여 검출을 행할 수 있다.
- [0024] 또한, 본 발명에 따른 검출 장치는, 각 인자 항목에 대해서, 대응하는 참조 데이터의 표준 편차의 평균값을 나타내는 참조 표준 편차, 및 인자 항목간의 상관관계를 나타내는 값의 평균값을 나타내는 참조 상관값을 산출하는 수단을 더 구비하고, 상기 검출 수단은, 상기 제1 지수가 상기 참조 표준 편차에 비해서 유의성을 가지고 증대하며, 상기 제2 지수가 상기 참조 상관값에 비해서 유의성을 가지고 감소하고, 또한, 상기 제3 지수가 상기 참조 표준 편차에 비해서 유의성을 가지고 증대한 경우에, 상기 클러스터에 포함되는 항목을 바이오마커의 후보로서 검출하는 것을 특징으로 한다.
- [0025] 따라서, 선출된 인자 항목이 바이오마커로서 적절한지의 여부를 판정할 수 있다.
- [0026] 또한, 본 발명에 따른 검출 장치에 있어서, 상기 검출 수단은, 클러스터에 포함되는 복수의 인자 항목의 유의성을, 측정 데이터의 통계값에 기초하여 검정하는 수단을 포함하고, 유의성이 있는 경우에, 클러스터에 포함되는

항목을 바이오마커의 후보로서 검출하는 것을 특징으로 한다.

[0027] 따라서, 검정에 의해 오류 검출의 발생을 극력 방지할 수 있다.

[0028] 또한, 본 발명에 따른 검출 장치는, 상기 복수의 인자 항목 중 어느 하나는, 유전자에 관한 측정 항목, 단백질에 관한 측정 항목, 대사물에 관한 측정 항목, 또는 생체로부터 얻어지는 화상에 관한 측정 항목인 것을 특징으로 한다.

[0029] 따라서, 유전자, 단백질, 대사물에 관해서 측정 항목을 인자 항목으로 함으로써, 생체 내의 생물 변화를 정량적으로 파악하여, 검출 결과의 신뢰성을 향상시킬 수 있다.

[0030] 또한, 본 발명에 따른 검출 방법은, 생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 검출 장치를 이용한 검출 방법으로서, 상기 검출 장치는, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 복수의 인자 항목을 복수의 클러스터로 분류하는 분류 단계와, 분류한 각 클러스터로부터, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화 및 각 인자 항목간의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 미리 설정된 선출 조건에 해당하는 클러스터를 선출하는 선출 단계와, 선출한 클러스터에 포함되는 인자 항목을 바이오마커의 후보로서 검출하는 검출 단계를 실행하는 것을 특징으로 한다.

[0031] 또한, 본 발명에 따른 검출 방법은, 생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 검출 방법으로서, 상이한 시간점에서 채취한 복수의 생체 샘플의 각각으로부터 측정된 하이 스루풋 데이터 중에서, 디퍼렌셜 생체 분자를 산출하는 분자 스크리닝 단계와, 상관성이 높은 생체 분자끼리를 하나의 클러스터로 하도록, 상기 분자 스크리닝 단계에서 선출한 상기 디퍼렌셜 생체 분자를 복수의 클러스터로 분류하는 클러스터화 단계와, 상기 클러스터화 단계에서 얻어진 복수의 클러스터 중에서, 클러스터 내의 생체 분자 간의 상관의 증대, 클러스터 내의 생체 분자의 표준 편차의 증대, 및 클러스터 내의 생체 분자와 다른 생체 분자 간의 상관의 저감이 가장 현저한 클러스터를, 상기 바이오마커의 후보로서 선취하는 후보 선택 단계와, 상기 후보 선택 단계에서 선출한 바이오마커의 후보가 상기 바이오마커인지의 여부를, 유의성 검정에 의해 판정하는 판정 단계를 실행하는 것을 특징으로 한다.

[0032] 상기 특징을 갖는 검출 방법은, 정상 상태에서부터 질병 상태로 천이하기 전의 질병 전 상태를 나타내는 경고 신호가 되는 바이오마커의 후보를 검출할 수 있다. 또한, 바이오마커를 특정할 수 있으면, 검출 대상으로부터 적은 수의 샘플을 채취하는 것만으로 질병 전 상태를 판별할 수 있다.

[0033] 또한, 본 발명에 따른 검출 프로그램은, 컴퓨터에, 생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터에 기초하여, 측정 대상인 생체의 증상의 지표가 되는 바이오마커의 후보를 검출하는 처리를 실행시키는 검출 프로그램으로서, 컴퓨터에, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 복수의 인자 항목을 복수의 클러스터로 분류하는 분류 단계와, 분류한 각 클러스터로부터, 상기 각 인자 항목의 각각의 측정 데이터의 시계열 변화 및 각 인자 항목간의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 미리 설정된 선출 조건에 해당하는 클러스터를 선출하는 선출 단계와, 선출한 클러스터에 포함되는 인자 항목을 바이오마커의 후보로서 검출하는 검출 단계를 실행시키는 것을 특징으로 한다.

[0034] 상기 특징을 갖는 검출 프로그램은, 컴퓨터에서 실행함으로써, 컴퓨터가, 본 발명에 따른 검출 장치로서 동작한다. 따라서, 정상 상태에서부터 질병 상태로 천이하기 전의 질병 전 상태를 나타내는 경고 신호가 되는 바이오마커의 후보를 검출할 수 있다. 또한, 바이오마커를 특정할 수 있으면, 검출 대상으로부터 적은 수의 샘플을 채취하는 것만으로 질병 전 상태를 판별할 수 있다.

발명의 효과

[0035] 본 발명에 따르면, 생체 샘플을 진단 대상으로부터 채취하고, 채취한 생체 샘플에 있어서, 질병 상태가 되기 직전의 질병 전 상태를 나타내는 경고 신호가 되는 바이오마커가 존재하는지의 여부를 조사함으로써, 진단 대상이 질병 전 상태에 있는지의 여부를 진단할 수 있다. 그 때문에, 질병 악화의 모델을 구축할 필요가 없어, 질병 악화의 구동 요소를 특정할 필요도 없으며, 질병이 되기 전의 단계에서, 질병의 조기 진단을 실현할 수 있다.

도면의 간단한 설명

[0036] 도 1은 질병의 진행 프로세스를 설명하는 모식도이다.

- 도 2는 본 발명의 검출 방법에 따른 DNB의 동적 특성을 예시하는 모식도이다.
- 도 3은 실시형태에 있어서의 DNB의 검출 방법의 일례를 나타내는 흐름도이다.
- 도 4는 실시형태에 있어서의 디퍼렌셜 생체 분자의 선출 처리의 일례를 나타내는 흐름도이다.
- 도 5는 실시형태에 있어서의 DNB의 후보의 선출 처리의 일례를 나타내는 흐름도이다.
- 도 6은 실시형태에 있어서의 DNB의 판정 처리의 일례를 나타내는 흐름도이다.
- 도 7은 실시형태에 있어서의 DNB에 의한 질병의 조기 진단 방법에 따른 진단 스케줄의 일례를 나타내는 도면이다.
- 도 8은 실시형태에 있어서의 DNB에 의한 질병의 조기 진단 방법의 일례를 나타내는 흐름도이다.
- 도 9는 종합 지수(I)에 비례한 질병 리스크를 표시하는 도형의 일례이다.
- 도 10은 종합 지수(I)에 비례한 질병 리스크를 표시하는 도형의 일례이다.
- 도 11은 본 발명에 따른 검출 장치의 구성예를 나타내는 블록도이다.
- 도 12는 본 발명에 따른 검출 장치에 의한 DNB의 검출 처리의 일례를 나타내는 흐름도이다.
- 도 13은 제1 검증예에 있어서의 진단용의 데이터를 나타내는 테이블이다.
- 도 14a는 제1 검증예에 있어서, 검출된 DNB의 후보의 표준 편차의 평균값의 시계열 변화의 예를 나타내는 그래프이다.
- 도 14b는 제1 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버간의 피어슨 상관 계수의 절대값의 평균값의 시계열 변화의 예를 나타내는 그래프이다.
- 도 14c는 제1 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버와 다른 유전자 간의 피어슨 상관 계수의 절대값의 평균값의 시계열 변화의 예를 나타내는 그래프이다.
- 도 14d는 제1 검증예에 있어서, 검출된 DNB의 후보의 종합 지수의 평균값의 시계열 변화의 예를 나타내는 그래프이다.
- 도 15는 제1 검증예에 있어서, 케이스 그룹의 유전자에 의해 구성된 네트워크에 있어서의 DNB의 동적 특성의 일례를 경시적으로 나타내는 맵이다.
- 도 16은 제2 검증예에 있어서의 진단 데이터의 테이블이다.
- 도 17a는 제2 검증예에 있어서, 검출된 DNB의 후보의 표준 편차의 평균값의 시계열 변화의 예를 나타내는 그래프이다.
- 도 17b는 제2 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버간의 피어슨 상관 계수의 절대값의 평균값의 시계열 변화의 예를 나타내는 그래프이다.
- 도 17c는 제2 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버와 다른 유전자 간의 피어슨 상관 계수의 절대값의 평균값의 시계열 변화의 예를 나타내는 그래프이다.
- 도 17d는 제2 검증예에 있어서, 검출된 DNB의 후보의 종합 지수의 평균값의 시계열 변화의 예를 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0037]

본 발명의 발명자들은, 하나의 샘플로부터 수천개의 정보, 즉 고차원 데이터를 얻는 것이 가능한 게놈 하이 스퀀럼 테크놀러지를 이용하여, 분기 프로세스 이론에 기초하여, 복잡 질병의 시간 발전의 수리 모델을 구축하고, 분자 네트워크 레벨에 있어서의 질병 악화의 진행 메커니즘의 연구를 행하였다. 그 결과, 질병 전 상태에 있어서, 임계 천이 발생 전의 직전 분기(돌연 악화) 상태를 검출할 수 있는 동적 네트워크 바이오마커(DNB: Dynamical Network Biomarker)의 존재를 해명하였다. 상기 동적 네트워크 바이오마커를 질병 전 상태의 경고 신호로서 이용하면, 질병 모델이 불필요하여, 적은 수의 샘플만으로 복잡 질병의 조기 진단을 실현할 수 있다. 이하에, 동적 네트워크 바이오마커에 기초한 본 발명을 실시하기 위한 형태에 대해서 설명한다.

- [0038] <이론 근거>
- [0039] 우선, 본 발명의 기초가 되는 이론 근거에 대해서 설명한다. 통상, 질병의 진행 프로세스에 관한 시스템[이하, 시스템(1)이라고 함]을 다음 식 (1)로 나타낼 수 있다.
- [0040] $Z(k+1)=f(Z(k);P)\cdots$ 식 (1)
- [0041] 여기서, $Z(k)=(z_1(k), \dots, z_n(k))$ 는, 시각(k)($k=0, 1, \dots$)에서 관측된 시스템(1)의 동적 상태를 나타내는 변수이며, 유전자 발현량, 단백질 발현량, 대사물 발현량 등의 정보로 할 수 있다. 상세하게는, 유전자, 단백질 등의 분자에 관한 농도, 개수 등의 정보이다. P는, 시스템(1)의 상태 천이를 구동하는 완만하게 변화하는 파라미터이며, 예컨대, SNP(Single Nucleotide Polymorphism), CNV(Copy Number Variation) 등의 유전적 요인, 메틸화, 아세틸화 등의 비유전적 요인 등의 정보로 할 수 있다. $f=(f_1, \dots, f_n)$ 는 $Z(k)$ 의 비선형 함수이다.
- [0042] 정상 상태, 질병 상태는 각각, 상태 방정식 $Z(k+1)=f(Z(k);P)$ 의 어트랙터로 나타낼 수 있다. 복잡 질병의 진행 프로세스는, 매우 복잡한 동적 특성을 갖기 때문에, 함수(f)는, 수천개의 변수를 갖는 비선형 함수이다. 더구나, 시스템(1)을 구동하는 요소(P)는 특정하기 어려운 것이다. 그 때문에, 정상 상태 및 질병 상태의 시스템 모델을 구축하여 해석하는 것은 대단히 곤란하다.
- [0043] 그에 대하여, 본 발명의 발명자들이 착안한 것은, 정상 상태에서부터 질병 상태로 천이하기 직전의 시스템의 임계 천이 상태, 즉 질병 전 상태이다. 일반적으로, 시스템(1)에는, 다음 특성을 갖는 평형점이 존재한다.
- [0044] 1. Z^* 는, 시스템(1)의 고정점이라고 하면, $Z^*=f(Z^*; P)$,
- [0045] 2. P_c 를 시스템이 분기하는 역치라고 하면, $P=P_c$ 일 때, 야코비 행렬 $\partial f(Z; P_c)/\partial Z|_{Z=Z^*}$ 의 하나의 실수 고유값 또는 한쌍의 복소 공역의 고유값의 절대값이 1이 된다,
- [0046] 3. $P \neq P_c$ 일 때, 일반적으로, 시스템(1)의 고유값의 절대값이 1이 아니다.
- [0047] 발명자들은, 상기 특성으로부터, 시스템(1)이 임계 천이 상태가 되면, 다음과 같은 특이한 특성이 나타나는 것을 이론적으로 해명하였다. 즉, 시스템(1)이 임계 천이 상태가 되면, 시스템(1)의 각 변수(z_1, \dots, z_n)의 각각을 노드로 하여 구성된 네트워크(1)에 있어서, 일부의 노드로 이루어지는 지배 그룹(서브 네트워크)이 나타난다. 임계 천이 상태에 나타나는 지배 그룹은, 이상적으로는, 다음과 같은 특이한 특성을 갖는다.
- [0048] (I) z_i 와 z_j 는 지배 그룹에 속하는 노드인 경우,
- [0049] $PCC(z_i, z_j) \rightarrow \pm 1$;
- [0050] $SD(z_i) \rightarrow \infty$;
- [0051] $SD(z_j) \rightarrow \infty$.
- [0052] (II) z_i 는 지배 그룹에 속하는 노드이지만, z_j 는 지배 그룹에 속하는 노드가 아닌 경우,
- [0053] $PCC(z_i, z_j) \rightarrow 0$;
- [0054] $SD(z_i) \rightarrow \infty$;
- [0055] $SD(z_j) \rightarrow$ 경계값.
- [0056] (III) z_i 와 z_j 는 지배 그룹에 속하는 노드가 아닌 경우
- [0057] $PCC(z_i, z_j) \rightarrow \alpha, \alpha \in (-1, 1) \setminus \{0\}$;
- [0058] $SD(z_i) \rightarrow$ 경계값;
- [0059] $SD(z_j) \rightarrow$ 경계값.
- [0060] 여기서, $PCC(z_i, z_j)$ 는 z_i 와 z_j 간의 피어슨 상관 계수이며, $SD(z_i), SD(z_j)$ 는 z_i 와 z_j 의 표준 편차이다.
- [0061] 즉, 네트워크(1)에 있어서, 상기 특이한 특성 (I)~(III)을 갖는 지배 그룹의 출현은, 시스템(1)이 임계 천이 상태(질병 전 상태)로 천이하고 있는 것의 징후로서 파악할 수 있다. 따라서, 상기 지배 그룹을 검출함으로써, 시스템(1)의 임계 천이를 검출할 수 있다. 즉, 상기 지배 그룹을 임계 천이, 즉, 질병 악화 직전의 질병 전 상태를 나타내는 경고 신호로 할 수 있다. 그렇게 하면, 시스템(1)이 아무리 복잡하여도, 구동 요소가 불명하여도, 경고 신호가 되는 지배 그룹만을 검출하면, 시스템(1)의 수리 모델을 직접 취급하는 일없이, 질병

전 상태인 것을 특정할 수 있다. 질병 전 상태인 것을 특정함으로써, 질병에 대한 사전 대책 및 조기 치료를 실현하는 것이 가능해진다.

[0062] 본 발명에 있어서, 질병 전 상태를 나타내는 경고 신호가 되는 상기 지배 그룹을 동적 네트워크 바이오마커(이하, DNB라고 약칭함)라고 칭한다.

[0063] <DNB의 특성 및 판정 조건>

[0064] 전술한 바와 같이, DNB는, 특이한 특성 (I)~(III)을 갖는 지배 그룹으로서, 복수의 노드로 이루어지는 서브 네트워크로서, 시스템(1)이 질병 전 상태가 될 때에, 네트워크(1)에 나타나는 것이다. 네트워크(1)에 있어서, 각 노드(z_1, \dots, z_n)를, 유전자, 단백질, 대사물 등의 생체 분자에 대해서 측정의 대상이 되는 인자 항목으로 하면, DNB는, 상기 특이한 특성 (I)~(III)을 만족시킨 일부의 생체 분자에 관한 인자 항목으로 이루어지는 그룹(서브 네트워크)이다.

[0065] 또한, 상기 특이한 특성 (I)~(III)에 기초하여, DNB를 판정하는 조건을 이하와 같이 정할 수 있다.

[0066] · 조건 (I): 네트워크(1)에 존재하는 일부의 생체 분자(유전자, 단백질 또는 대사물)로 이루어지는 그룹으로서, 상기 그룹 내의 생체 분자 간의 피어슨 상관 계수 PCC의 절대값의 평균값이 현저히 증대한다.

[0067] · 조건 (II): 상기 그룹 내의 생체 분자와 다른 생체 분자 간의 피어슨 상관 계수 OPCC의 절대값의 평균값이 현저히 저감한다.

[0068] · 조건 (III): 상기 그룹 내의 생체 분자의 표준 편차(SD)의 평균값이 현저히 증가한다.

[0069] 상기 DNB를 판정하는 조건 (I)~(III)을 동시에 만족시킨 생체 분자로 이루어지는 그룹을 DNB로 인정한다.

[0070] DNB의 특성을 직관적으로 설명하기 위해, 다음에, 6개의 노드를 갖는 네트워크를 예로 하여, 상기 네트워크에 있어서의 DNB의 동적 특성을 설명한다. 도 2는 본 발명의 검출 방법에 따른 DNB의 동적 특성을 예시하는 모식도이다. 도 2의 a는 질병의 진행 프로세스로서, 정상 상태, 질병 전 상태 및 질병 상태를 나타내고 있다. 도 2의 b, c 및 d는 진행 프로세스의 과정에 대해, 정상 상태, 질병 전 상태 및 질병 상태에 있어서, 모델화한 시스템 (1)의 안정성을, 포텐셜 함수로서 나타내고, 횡축에 경과를 나타내는 시각을 들며, 종축에 포텐셜 함수의 값을 들어 개념적으로 나타낸 그래프이다. 또한, 도 2의 e, f 및 g는 정상 상태, 질병 전 상태 및 질병 상태의 각각에 대응한 시스템(1)의 네트워크 상태의 일례를 개념적으로 나타내고 있다. 또한, 도 2의 h는 질병 전 상태에 있어서의 DNB의 인자 항목이 되는 분자 농도의 시간 변화의 일례를 나타내고 있다.

[0071] 노드($z_1 \sim z_6$)는, 유전자, 단백질, 대사물 등의 상이한 종류의 생체 분자에 관한 인자 항목을 나타내는 것이며, 노드($z_1 \sim z_6$) 간의 접속 라인은, 노드 간의 상관을 나타내고, 라인의 굵기는, 피어슨 상관 계수(PCC)의 크기를 나타내며, $z_1 \sim z_6$ 을 둘러싸는 동그라미표의 묘화 방법은, 각 노드의 표준 편차(SD)의 크기를 나타내고 있다. 접속 라인의 묘화 방법으로서, 동그라미표 내가 공백인 경우에 표준 편차(SD)는 가장 작고, 1방향의 사선, 2방향의 사선이 뒀에 따라 표준 편차(SD)가 큰 것을 나타내고 있다.

[0072] 정상 상태에 있어서, 도 2의 e에 나타내는 바와 같이, 각 노드는, 상호의 상관 및 각각의 표준 편차가 모두 균등하며, 중례벨로 되어 있다. 그러나, 질병 전 상태가 되면, 다른 노드에 비해서, 현저한 특이한 특성을 나타내는 그룹($z_1 \sim z_3$)이 나타난다. 상기 그룹 내의 노드($z_1 \sim z_3$)는, 도 2의 f에 나타내는 바와 같이, 상호간의 피어슨 상관 계수가 현저히 증대하며, 다른 노드($z_4 \sim z_6$)와의 사이의 피어슨 상관 계수가 현저히 저하하고 있다. 또한, 상기 그룹 내의 노드($z_1 \sim z_3$)의 표준 편차는 커져 있다. 그 이유는, 상기 그룹 내의 노드($z_1 \sim z_3$)는, 도 2의 h에 나타내는 바와 같이, 다른 시각($t=1, t=2, t=3$)에 있어서의 농도의 변화가 격심하기 때문이다.

[0073] 그러나, 질병 상태로 천이한 후는, 도 2의 g에 나타내는 바와 같이, 상기 그룹 내의 각 노드($z_1 \sim z_3$)의 표준 편차는 조금 커져 있지만, 각 노드 간의 피어슨 상관 계수는 균등하게 중례벨로 되돌아가 있다. 즉, 상기 그룹($z_1 \sim z_3$)의 상기 특이한 특성이 소실되어 있다.

[0074] 도 2에 개념적으로 나타내는 바와 같이, 질병 전 상태에 있어서, 그 특성을 갖는 일부의 노드를 포함하는 지배 그룹이 출현한다. DNB라고 칭하는 이러한 지배 그룹의 출현은, 질병 전 상태에 있는 것을 나타내며, 질병 상태로의 천이를 예견할 수 있는 경고 신호이고, 질병의 조기 진단을 위한 바이오마커로서 사용할 수 있다. 또한, DNB는, 통상의 질병 진단에 사용되고 있는 정적 바이오마커와 상이하며, 특성이 변화하고 있는 네트워크에 있어서 나타나는 서브 네트워크이다. 그래서, 본원에 있어서, 상기 지배 그룹을 DNB(동적 네트워크 바이오마커)라고 부르는 것으로 하였다.

- [0075] <경고 신호>
- [0076] 전술한 바와 같이, DNB는, 질병 전 상태를 나타내는 경고 신호로서, 질병의 조기 진단에 사용할 수 있다. 상기 경고 신호의 강도를 측정하는 것으로서, 전술한 DNB 내의 노드 간의 피어슨 상관 계수(PCC)의 절대값의 평균값, DNB 내의 노드와 다른 노드 간의 피어슨 상관 계수(OPCC)의 절대값의 평균값, 및 DNB의 표준 편차(SD)를 이용할 수 있다. 또한, DNB의 특성을 종합적으로 반영할 수 있는 종합 지수(I)를 도입할 수 있다. 본 발명에 있어서, 일례로서, 다음 식 (2)로 나타내는 종합 지수(I)를 도입한다.
- [0077] $I = SDd \times PCCd / OPCCd \dots$ 식 (2)
- [0078] 식 (2)에 있어서, PCCd는 DNB 내의 노드 간의 피어슨 상관 함수의 절대값의 평균값을 나타내고, OPCCd는 DNB 내의 노드와 다른 노드의 피어슨 상관 계수의 절대값의 평균값을 나타내며, SDd는 DNB 내의 노드의 표준 편차의 평균값을 나타낸다. 상기 식 (2)로부터 알 수 있듯이, SDd 및 PCCd가 증대하며, 또한, OPCCd가 감소하면, 상기 종합 지수(I)는 대폭 증대한다. 그 때문에, DNB의 특성을 높은 감도로 검출할 수 있다. 또한, 상기 종합 지수(I)의 값으로부터, 질병 상태와의 거리를 어느 정도 파악할 수 있다.
- [0079] <DNB의 검출 방법>
- [0080] 도 3은 실시형태에 있어서의 DNB의 검출 방법의 일례를 나타내는 흐름도이다. 본 발명의 검출 방법에 있어서는, 우선, 생체에 관한 측정에 의해 측정 데이터를 얻는 것이 필요하다. DNA 칩 등의 하이 스루풋 기술을 이용하면, 하나의 생체 샘플로부터 2만개 이상의 유전자를 측정할 수 있다. 통계적으로 분석하기 위해, 본 발명에 있어서, 측정 대상으로부터 복수(6개 이상)의 생체 샘플을 상이한 시간점에서 채취하고, 채취한 생체 샘플을 측정하여 얻어진 측정 데이터를 집계하여 통계 데이터를 취득한다. 본 발명에 있어서의 DNB의 검출 방법은, 주로, 도 3에 나타내는 바와 같이, 하이 스루풋 데이터의 취득 처리(s1)와, 디퍼렌셜 생체 분자의 선출 처리(s2)와, 클러스터화 처리(s3)와, DNB의 후보의 선출 처리(s4)와, 유의성 분석에 의한 DNB의 판정 처리(s5)를 포함한다. 다음에, 이들 각 처리에 대해서 상세하게 설명한다.
- [0081] 단계 s1의 하이 스루풋 데이터의 취득 처리는, 검출 대상이 되는 샘플을 케이스 샘플로 하고, 참조용 샘플을 컨트롤 샘플로 하여, 각 샘플로부터, 각각의 하이 스루풋 기술을 이용한 생리학적 데이터, 즉, 생체 분자의 발현량의 측정 데이터(예컨대, 마이크로 어레이 데이터)를 취득하는 처리이다. 참조용 샘플이란, 검사의 대상이 되는 환자로부터 사전에 채취한 샘플, 채취 시에 최초로 채취한 샘플 등의 샘플이며, 측정 장치의 교정 등의 목적을 갖는 컨트롤 샘플로서 이용된다. 컨트롤 샘플은, 반드시 필요하지 않지만, 오차 요인을 배제하고, 측정의 신뢰성을 향상시키기 때문에 유용하다.
- [0082] 단계 s2의 디퍼렌셜 생체 분자의 선출 처리는, 발현량에 현저한 변화를 나타낸 생체 분자를 선택하는 처리이다. 도 4는 실시형태에 있어서의 디퍼렌셜 생체 분자의 선출 처리의 일례를 나타내는 흐름도이다. 도 4는 도 3에 나타낸 단계 s2의 디퍼렌셜 생체 분자의 생체 처리를 상세하게 나타낸 것이다.
- [0083] 도 4에 나타내는 바와 같이, 우선, n개의 케이스 샘플로부터 각각 측정된 하이 스루풋 데이터(생체 분자의 발현량)에 기초한 통계 데이터를 D1c로 하고, 컨트롤 샘플로부터 측정된 데이터를 Dr로 한다(s21). 다음에, 각 케이스 샘플 생체 분자(D1c)에 대하여, 스튜던트의 t 검정을 행하고, 컨트롤 샘플의 하이 스루풋 데이터(Dr)에 비해서 발현량에 현저한 변화를 나타낸 생체 분자(D2c)를 선출한다(s22). 단계 s22에서는, 발현량에 현저한 변화를 나타낸 생체 분자(D2c)를 선출하는 방법으로서, 스튜던트의 t 검정을 예시하고 있지만, 특별히 방법을 한정하는 것이 아니다. 예컨대, 만·휘트니의 U 검정 등의 다른 검정 방법을 적용하는 것도 가능하고, 이러한 논파라메트릭 방법에 따른 검정은, 모집단이 되는 D1c가 정규 분포에 따르지 않는 경우에, 특히 유효하다. 또한, 스튜던트의 t 검정을 행하는 경우에 있어서도, 유의 수준(α)의 값은, 0.05, 0.01 등의 값을 적절하게 설정할 수 있다.
- [0084] 다음에, 오발현율[FDR(False Discovery Rate)]을 이용하여, 각 케이스 샘플 생체 분자(D2c)에 대하여, 다중 비교(multiple comparisons) 또는 복수의 스튜던트의 t 검정을 보정하고, 보정 후의 각 케이스 샘플 유전자 또는 단백질 데이터(D3c)를 선출한다(s23). 다음에, 2배 변경 방법(The two-fold change method)을 이용하여, 보정 후의 각 케이스 샘플 유전자 또는 단백질 데이터(D3c)로부터 표준 편차(SD)가 비교적 현저히 변화하는 Dc를 디퍼렌셜 생체 분자로서 선출한다. 여기서 선출된 디퍼렌셜 생체 분자(Dc)는, 컨트롤 샘플의 생체 분자(Dr)에 비해서, 현저한 차이를 나타낼 뿐만 아니라, 자기의 평균값으로부터도 크게 이탈하고 있다. 단계 s23에 있어서도, 검정 방법은 스튜던트의 t 검정에 한정되는 것이 아니다.
- [0085] 다음에, 클러스터화 처리(도 3의 s3)를 행한다. 여기서 말하는 클러스터화 처리란, 복수의 생체 분자를, 서로

상관이 높은 그룹으로 각각 분류하는 처리이며, 생체 분자가 분류된 각각의 그룹을 클러스터라고 칭한다. 즉, 서로 상관이 높은 생체 분자를 하나의 클러스터로 하도록, 도 4에 나타내는 단계 s24에서 선택한 디퍼렌셜 생체 분자(Dc)를, n개의 클러스터로 분류한다. 얻어진 모든 클러스터는, 잠재적인 지배 그룹, 즉, 검출하여야 하는 DNB의 후보가 된다.

[0086] 다음에, 도 3에 나타내는 DNB의 후보의 선출 처리(s4)를 행한다. 도 5는 실시형태에 있어서의 DNB의 후보의 선출 처리의 일례를 나타내는 흐름도이다. 도 5는 도 3에 나타낸 단계 s4의 DNB의 후보의 선출 처리를 상세하게 나타낸 것이다. 즉, DNB의 후보의 선출 처리는, 도 5에 나타내는 DNB의 후보의 선출 처리의 흐름도에 기초하여 처리를 행한다. 도 5에 나타내는 순환 루프로, 모든 클러스터(k)(k=1, ..., n)에 대해서, 클러스터 내의 노드 간의 피어슨 상관 계수의 절대값의 평균값[PCCd(k)], 클러스터 내의 노드와 다른 노드 간의 피어슨 상관 함수의 절대값의 평균값[OPCCd(k)], 클러스터 내의 노드의 표준 편차의 평균값[SDd(k)], 및 종합 지수[I(k)]를 계산한다(s41~s46). 그리고, 모든 클러스터로부터, 종합 지수(I)의 값이 가장 큰 클러스터를 DNB의 후보로서 선출한다(s47).

[0087] 다음에, 도 3에 나타내는 유의성 분석에 의한 DNB의 판정 처리(도 3의 s5)를 행한다. 도 6은 실시형태에 있어서의 DNB의 판정 처리의 일례를 나타내는 흐름도이다. 도 6은 도 3에 나타낸 단계 s5의 우위성 분석에 의한 DNB의 판정 처리를 상세하게 나타낸 것이다. 즉, 진술한 DNB의 판정 조건 (I)~(III)에 기초하여, 단계 s47에서 DNB의 후보로서 선출된 클러스터(m)는 DNB인지의 여부를 판정한다. 여러가지 유의성 분석에 의해 판정할 수 있지만, 일례로서, 도 6에 나타내는 DNB의 판정 처리의 흐름도에 기초하여 처리를 행한다.

[0088] 도 6에 나타내는 바와 같이, 우선, 컨트롤 샘플로부터 취득한 데이터의 각 노드 간의 피어슨 상관 계수의 절대값의 평균값(PCCdr), 및 각 노드의 표준 편차의 평균값(SDdr)을 각각 계산한다(s51, s52). 그리고, 컨트롤 샘플의 피어슨 상관 계수의 평균값(PCCdr)에 비해서, 단계 s47에서 선출한 클러스터(m) 내의 각 노드 간의 피어슨 상관 계수의 절대값의 평균값[PCCd(m)]이 유의하게 증대하였는지의 여부를 판단한다(s53). 유의하게 증대하고 있지 않다고 판단한 경우(NO), DNB가 존재하지 않는다고 하는 결과를 출력하고, 처리를 종료한다(s57). 한편, 유의하게 증대하였다고 판단한 경우(YES), 다음 단계 s54로 진행한다. 단계 s54에 있어서, 컨트롤 샘플의 피어슨 상관 계수의 평균값(PCCdr)에 비해서, 클러스터(m) 내의 각 노드와 다른 노드의 피어슨 상관 계수의 평균값[OPCCd(m)]이 유의하게 저감하였는지의 여부를 판단한다(s54). 유의하게 저감하고 있지 않다고 판단한 경우(NO), DNB가 존재하지 않는다고 하는 결과를 출력하고(s57), 처리를 종료한다. 한편, 유의하게 저감하였다고 판단한 경우(YES), 다음 단계 s55로 진행한다. 단계 s55에 있어서, 컨트롤 샘플의 표준 편차 평균값(SDr)에 비해서, 클러스터(m) 내의 노드의 표준 편차 평균값[SDd(m)]이 유의하게 증대하였는지의 여부를 판단한다. 유의하게 증대하고 있지 않다고 판단한 경우(NO), DNB가 존재하지 않는다고 판단하고(s57), 처리를 종료시킨다. 한편, 유의하게 증대하였다고 판단한 경우, 클러스터(m)를 DNB로 인정하고(s56), 처리를 종료시킨다.

[0089] <DNB에 의한 질병의 조기 진단 방법>

[0090] 진단 스케줄로서는, 일정한 간격을 두고, 복수회 진단을 행하고, 매회의 진단에서 수개의 샘플을 취하는 것이 바람직하다. 도 7은 실시형태에 있어서의 DNB를 이용한 질병의 조기 진단의 진단 스케줄의 일례를 나타내는 도면이다. 도 7에 나타내는 바와 같이, 복수의 단계(단계-1, 단계-2, ..., 단계-T)에서, 샘플을 채취한다. 각 단계에서 채취하는 샘플의 수는, 데이터의 정밀도를 확보하기 위해, 통상, 하나의 단계에 있어서, 6개 이상의 샘플을 채취하는 것이 바람직하다. 2개의 연속한 단계 사이의 간격은, 질병의 상황에 따라, 수일, 수주간, 수개월, 또는 몇년으로 길게 설정할 수 있지만, 각 단계에서는, 짧은 기간에 있어서 상이한 시간점에서 샘플을 채취하는 것이 바람직하다. 예컨대, 6개의 샘플을 1일 중의 6개의 시간점에서 채취한다. 각 시간점 사이의 간격은, 상황에 따라, 예컨대, 수분간, 수시간으로 할 수 있다.

[0091] 도 8은 실시형태에 있어서의 DNB에 의한 질병의 조기 진단 방법의 일례를 나타내는 흐름도이다. 도 8에 나타내는 바와 같이, DNB에 의한 질병의 조기 진단 방법은, 주로, 샘플 채취 처리(s100)와, 디퍼렌셜 생체 분자의 선출(s200)과, DNB의 후보의 선출 처리(s300)와, 유의성 분석에 의한 DNB의 판정 처리(s400)와, 진단 결과의 출력 처리(s500)를 포함한다. 다음에, 이들 처리 내용에 대해서 상세하게 설명한다.

[0092] 샘플 채취 처리(s100): 일반의 질병 진단 방법과 동일하게, 진단하는 질병에 따라, 필요한 생리 데이터를 취하기 위한 샘플을 채취한다. 예컨대, 간장 장애의 경우, 혈액, 간장의 조직 등의 샘플을 채취한다.

[0093] 또한, 진단할 때에 진단 대상으로부터 채취한 샘플을 케이스 샘플로 하는 것 외에, 참조용의 샘플로서, 진단 대상 이외의 건강한 사람으로부터 채취한 샘플, 또는 진단 대상으로부터 최초로 채취한 샘플을 컨트롤 샘플로 할

수 있다.

- [0094] 디퍼렌셜 생체 분자의 선출(s200): 도 4에 나타낸 디퍼렌셜 생체 분자의 선출 처리 흐름도에 따라, 상기 단계 s100에서 채취한 샘플로부터, 디퍼렌셜 생체 분자를 선출한다.
- [0095] DNB의 후보의 선출 처리(s300): 도 5에 나타낸 DNB의 후보의 선출 흐름도에 따라, 상기 단계 s200에서 선출한 디퍼렌셜 생체 분자로부터, DNB의 후보가 되는 지배 그룹을 선출한다.
- [0096] 유의성 분석에 의한 DNB의 판정 처리(s400): 도 6에 나타낸 유의성 분석에 의한 DNB의 판정 방법을 나타내는 흐름도에 따라, 상기 단계 s300에서 선출한 DNB의 후보가 DNB인지의 여부를 판정한다.
- [0097] 진단 결과의 출력(s500): 상기 단계 s400에 있어서, DNB가 존재하지 않는다고 판단한 경우, 상기 단계 s300에서 선출한 DNB의 후보의 데이터를 다음번의 진단의 참조 데이터로서 기억 장치에 보존하고, 이상이 확인되지 않은 취지의 진단 결과를 출력한다. 한편, 상기 단계 s400에 있어서, DNB로 인정된 클러스터가 존재한다고 판단한 경우, 인정된 클러스터의 생체 분자 데이터를 DNB의 멤버로서 기억하고, 질병 전 상태에 있는 취지의 진단 결과를 출력한다. 또한, 검출한 DNB에 관련지은 진단 결과를 출력할 수도 있다. 또한, 여기서는, 진단 결과로서 나타내고 있지만, 의사에 의한 진단의 참고가 되는 결과이다. 즉, 단계 s500에서 출력되는 진단 결과는, 의사에 의한 진단 그 자체가 아니라, 의사의 진단의 보조로서, 진단을 지원하도록 진단의 참고가 되는 출력 데이터이다.
- [0098] 예컨대, 진단 결과로서 DNB의 특성을 종합적으로 반영하는 종합 지수(I)를 출력할 수 있다. 상기 종합 지수(I)가 크면 클수록, 분기점에 가깝기 때문에, 종합 지수(I)에 비례한 질병 리스크를 직관적으로 볼 수 있는 도형 또는 화상의 형태로 출력하면, 워닝 효과가 높아진다.
- [0099] 도 9 및 도 10은 종합 지수(I)에 비례한 질병 리스크를 표시하는 도형의 일례이다. 도 9에 있어서, 화살표 전체는 질병 전 상태(발병 전기)를 나타내고, 화살표에 의해 나타내는 방향의 흐름이 질병 상태(발병)의 경시 변화를 가리키고 있으며, 화살표 내의 좌측에 위치하는 마름모꼴 마크는 진단에서 얻어진 종합 지수(I)의 값에 따라 위치가 변화하는 발병 리스크 포인터이다. 종합 지수(I)의 값이 크면 클수록 마름모꼴 마크가 화살표의 우측에 근접하도록 설정한다.
- [0100] 또한, 이전에도 DNB를 이용하여 질병의 조기 진단을 받은 적이 있는 경우, 전회 진단에서 얻어진 종합 지수와 함께, 도 10과 같이, 종합 지수(I)에 비례한 질병 리스크를 표시할 수 있다. 도 10에 있어서, 점선의 마름모꼴 마크가 2011년 7월 1일의 진단에서 얻어진 종합 지수를 나타내고, 실선의 마름모꼴 마크가 2011년 9월 1일의 진단에서 얻어진 종합 지수를 나타내고 있다. 마름모꼴 마크의 위치 변동으로부터, 질병 상태에 근접하고 있는 것을 직관적으로 판단할 수 있다.
- [0101] 또한, DNB에 관련지은 정보로서, 검출한 DNB를 포함한 전체 네트워크, 또는 상기 DNB를 포함한 일부의 네트워크의 맵(예컨대, 후술하는 도 15)을 출력할 수 있다.
- [0102] 또한, DNB의 멤버인 생체 분자의 리스트를 출력할 수도 있다. 전술한 바와 같이, DNB는, 정상 상태에서부터 질병 상태로 전이하는 질병 전 상태로 나타나는 것이지만, DNB로서 검출된 생체 분자, 즉 유전자, 단백질 또는 대사물 자신이, 반드시 질병을 악화시키는 요인이 되는 병적 유전자, 단백질 또는 대사물이라고는 한정되지 않는다. 단, DNB 멤버의 일부가 질병과 관련하고 있는 것이 판명되어 있다.
- [0103] 그 때문에, 검출한 DNB의 멤버에 포함된, 특정 질병과 관련성이 있는 것(유전자, 단백질 또는 대사물)을 추출하면, 예컨대, 의사의 진단에 의해, 검진자 등의 진단 대상이 발증할 가능성이 있는 질병을 어느 정도 파악할 수 있다.
- [0104] 따라서, 전술한 「진단 결과의 출력」(도 8의 s500)에 이어서, 유전자, 단백질 또는 대사물과 질병의 대응 관계를 기억한 데이터 베이스를 이용하여, 검출한 DNB로부터 질병과 관련성이 있는 유전자, 단백질 또는 대사물을 추출하여, 진단의 참고가 되는 진단 결과로서 출력할 수 있다.
- [0105] 그렇게 하면, 건강 진단을 받는 사람으로부터 예컨대 혈액을 채취하여, 그 혈액으로부터 얻은 유전자, 단백질 또는 대사물의 데이터로부터 DNB를 검출한 경우, 상기 DNB에 포함되어 있는 유전자, 단백질 또는 대사물과 관련성이 있는 질병을 어느 정도 특정할 수 있어, 검진자가 갖는 잠재적 질병을 조기 단계에서 진단할 수 있다.
- [0106] <검출 장치>
- [0107] 이상 상세하게 서술한 DNB의 검출 방법은, 컴퓨터를 이용한 검출 장치로서 본 발명을 구현화할 수 있다. 도 11은 본 발명에 따른 검출 장치의 구성예를 나타내는 블록도이다. 도 11에 나타내는 검출 장치(1)는, 퍼스널 컴퓨

터, 서버 컴퓨터에 접속되는 클라이언트 컴퓨터, 그 외 각종 컴퓨터를 이용하여 실현된다. 검출 장치(1)는, 제어부(10), 기록부(11), 기억부(12), 입력부(13), 출력부(14), 취득부(15) 등의 각종 기구를 구비하고 있다.

[0108] 제어부(10)는, CPU(Central Processing Unit) 등의 회로를 이용하여 구성되며, 검출 장치(1) 전체를 제어하는 기구이다.

[0109] 기록부(11)는, HDD(Hard Disk Drive) 등의 자기 기록 기구, SSD(Solid State Disk) 등의 불휘발성 반도체 기록 기구 등의 불휘발성의 보조 기록 기구이다. 기록부(11)에는, 본 발명에 따른 검출 프로그램(11a) 등의 각종 프로그램 및 데이터가 기록되어 있다.

[0110] 기억부(12)는, SDRAM(Synchronous Dynamic Random Access Memory), SRAM(Static Random Access Memory) 등의 휘발성의 주기억 기구이다.

[0111] 입력부(13)는, 키보드, 마우스 등의 하드웨어, 및 드라이버 등의 소프트웨어를 구비하는 입력 기구이다.

[0112] 출력부(14)는, 모니터, 프린터 등의 하드웨어, 및 드라이버 등의 소프트웨어를 구비하는 출력 기구이다.

[0113] 취득부(15)는, 외부로부터 각종 데이터를 취득하는 기구이다. 구체적으로는, 통신망을 통해 데이터를 취득하는 LAN(Local Area Network) 포트, 측정 기기에 접속 가능한 병렬 케이블 등의 전용선에 접속하는 포트 등의 각종 하드웨어, 및 드라이버 등의 소프트웨어이다.

[0114] 그리고, 기록부(11)에 기록되어 있는 검출 프로그램(11a)을, 기억부(12)에 기억하고, 제어부(10)의 제어에 기초하여 실행함으로써, 컴퓨터는, 검출 프로그램(11a)에 관한 각종 순서를 실행하여, 본 발명의 검출 장치(1)로서 기능한다. 또한, 편의 상, 기록부(11) 및 기억부(12)로서 나누고 있지만, 양자 모두 각종 정보의 기록이라고 하는 동일한 기능을 갖는 것이며, 장치의 사양, 운용 형태 등에 따라 어느 기구에 기록시킬지는, 적절하게 결정하는 것이 가능하다.

[0115] 도 12는 본 발명에 따른 검출 장치(1)에 의한 DNB의 검출 처리의 일례를 나타내는 흐름도이다. 본 발명에 따른 검출 장치(1)의 처리는, 전술한 DNB의 검출 처리를 실행하는 것이다. 검출 장치(1)의 제어부(10)는, 취득부(15)에 의해, 생체에 관한 측정에 의해 얻어진 복수의 인자 항목에 대한 측정 데이터를 취득한다(Sc1). 단계 Sc1은, 도 3의 단계 s1로서 나타낸 하이 스트루프 데이터의 취득 처리에 대응한다. 또한, 여기서는 컴퓨터 처리의 대상으로서 표현하기 위해, 인자 항목으로서 표현하고 있지만, 여기서 말하는 인자 항목이란, 전술한 DNB의 노드가 될 수 있는 유전자에 관한 측정 항목, 단백질에 관한 측정 항목, 대사에 관한 측정 항목 등의 측정 항목을 나타내고 있다. 또한, CT 스캔 등의 측정 장치에서 출력되는 체내 화상으로부터 얻어지는 화상에 관한 측정 항목을 이용하는 것도 가능하다.

[0116] 제어부(10)는, 취득한 인자 항목의 각각의 측정 데이터가, 유의성을 가지고 경시적으로 변화하고 있는지의 여부를 검정하고, 검정한 결과에 기초하여 디퍼렌셜 생체 분자를 선출한다(Sc2). 단계 Sc2는, 도 3의 단계 s2로서 나타낸 디퍼렌셜 생체 분자의 선출 처리에 대응한다.

[0117] 따라서, 단계 Sc2의 처리는, 제어부(10)가, 각 인자 항목의 측정 데이터, 및 인자 항목 및 시계열마다 미리 설정되어 있는 참조 데이터와의 비교 결과에 기초하여, 유의성에 관한 검정을 행하고(Sc21), 경시 변화에 유의성이 있다고 검정된 인자 항목을 선출하는 처리(Sc22)를 포함한다. 즉, 도 4에 나타낸 각종 처리를 실행한다. 또한, 검출 장치(1)가 참조 데이터로서 처리하는 데이터란, 컨트롤 샘플이며, 예컨대, 최초로 취득한 샘플을 컨트롤 샘플로 하는 등의 설정에 기초하여, 검출 장치(1)는, 상기 샘플에 대하여 참조 데이터로서의 취급을 행한다.

[0118] 제어부(10)는, 선출한 각 인자 항목의 각각의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 복수의 인자 항목을 복수의 클러스터로 분류한다(Sc3). 단계 Sc3은, 도 3의 단계 s3으로서 나타낸 클러스터화 처리에 대응한다.

[0119] 제어부(10)는, 분류한 각 클러스터로부터, 각 인자 항목의 각각의 측정 데이터의 시계열 변화 및 각 인자 항목 간의 측정 데이터의 시계열 변화의 상관 관계에 기초하여 미리 설정된 선출 조건에 해당하는 클러스터를 선출한다(Sc4). 단계 Sc4는, 도 3의 단계 s4로서 나타낸 DNB의 후보의 선출 처리에 대응한다.

[0120] 따라서, 단계 Sc4의 처리는, 제어부(10)가, 각각의 클러스터에 대하여, 클러스터 내의 각 인자 항목의 측정 데이터의 상관을 나타내는 값의 평균값을 제1 지수로서 산출하고(Sc41), 클러스터 내의 인자 항목의 측정 데이터와 상기 클러스터 외의 인자 항목의 측정 데이터 간의 상관을 나타내는 값의 평균값을 제2 지수로서 산출하며(Sc42), 클러스터 내의 각 인자 항목에 대해서 측정 데이터의 표준 편차의 평균값을 제3 지수로서 산출하는 처

리(Sc43)를 포함한다. 또한, 단계 Sc4의 처리는, 제어부(10)가, 제1 지수와, 제2 지수와, 제3 지수의 역수와의 곱에 기초한 종합 지수를 산출하고(Sc44), 산출한 종합 지수가 최대인 클러스터를 선출하는 처리(Sc45)를 포함한다. 즉, 도 5에 나타낸 각종 처리를 실행한다. 또한, 제1 지수, 제2 지수, 제3 지수 및 종합 지수로서는, 예컨대, 클러스터 내의 노드 간의 피어슨 상관 계수의 절대값의 평균값[PCCd(k)], 클러스터 내의 노드와 다른 노드 간의 피어슨 상관 함수의 절대값의 평균값[OPPCd(k)], 클러스터 내의 노드의 표준 편차의 평균값[SDd(k)], 및 종합 지수[I(k)]가 각각 이용된다.

[0121] 제어부(10)는, 선출한 클러스터에 포함되는 인자 항목을 바이오마커의 후보로서 검출한다(Sc5). 단계 Sc5는, 도 3의 단계 s5로서 나타낸 DNB의 판정 처리에 대응한다.

[0122] 따라서, 단계 Sc5의 처리는, 제어부(10)가, 각 인자 항목에 대해서, 대응하는 참조 데이터의 표준 편차의 평균값을 나타내는 참조 표준 편차를 산출하고(Sc51), 인자 항목간의 상관을 나타내는 값의 평균값을 나타내는 참조 상관값을 산출한다(Sc52). 그리고, 제1 지수가 참조 표준 편차에 비해서 유의성을 가지고 증대하며, 제2 지수가 참조 상관값에 비해서 유의성을 가지고 감소하고, 또한, 제3 지수가 참조 표준 편차에 비해서 유의성을 가지고 증대한 경우에, 상기 클러스터에 포함되는 항목을 바이오마커로서 검출하는 처리(Sc53)를 포함한다. 즉, 도 6에 나타낸 각종 처리를 실행한다. 또한, 참조 표준 편차 및 참조 상관값으로서, 예컨대, 각 노드 간의 피어슨 상관 계수의 절대값의 평균값(PCCdr) 및 각 노드의 표준 편차의 평균값(SDdr)이 각각 이용된다.

[0123] 그리고, 제어부(10)는, 바이오마커의 후보로서 검출한 인자 항목을 출력부(14)로부터 출력하고(Sc6), 처리를 종료한다.

[0124] <제1 검증예>

[0125] 본 발명의 DNB에 의한 질병의 조기 진단 방법의 진단 정밀도를 검증하기 위해, 폐 장애가 야기된 마우스의 실험 데이터를 이용하여, 본 발명의 진단 방법으로 진단을 행하고, 그 진단 결과를 실제의 병상 진행 상황과 비교하여, 본 발명의 진단 방법의 유효성을 검증하였다. 다음에, 상기 검증예를 상세하게 설명한다. 실험 데이터는, 복수의 실험용 CD-1 수컷 마우스를 케이스 그룹과, 컨트롤 그룹으로 나누어, 케이스 그룹을 통상의 공기 환경에, 컨트롤 그룹을 유독 가스인 포스겐이 포함되어 있는 공기 환경에 두며, 그리고, 양 그룹의 마우스의 건강 상황을 관찰하여, 포스겐의 흡입에 의한 급성 폐 손상의 분자 레벨의 메커니즘을 조사한다고 하는 실험으로부터 얻어진 것이다. 상기 실험 데이터를 이용하여, 본 발명의 진단 방법을 이용하여, 포스겐 폭로되어 있는 케이스 그룹의 마우스의 건강 상황을 진단하였다. 통상, 마우스는 일정량의 포스겐을 흡입하면, 포스겐 유발성 폐 장애를 발증한다.

[0126] 도 13은 제1 검증예에 있어서의 진단용의 데이터를 나타내는 테이블이다. 도 13에 나타내는 바와 같이, 진단 대상은, 포스겐 유발성 폐 장애에 걸린 마우스(CD-1 수컷 마우스)이며, 샘플 채취 대상은, 진단 대상이 되는 케이스 그룹 및 참조 대상이 되는 컨트롤 그룹의 마우스의 폐 조직이고, 샘플링 포인트는, 실험 개시 후의 0, 0.5, 1, 4, 8, 12, 24, 48, 72시간이 경과한 시간점이며, DNB의 검출에 이용한 유전자의 수는 22690개이다.

[0127] 본 발명의 진단 방법을 이용하여, 구체적으로, 다음 처리를 행하였다.

[0128] 우선, 각각의 샘플로부터 측정된 하이 스루풋 유전자 데이터로부터 디퍼렌셜발현 유전자를 선택한다. 각각의 샘플링 포인트에 있어서, 6개의 케이스 샘플과 6개 컨트롤 샘플이 제공되어 있다. 최초의 샘플링 포인트(0 h)에 있어서, 케이스 샘플의 데이터는 컨트롤 샘플의 데이터와 동일하다.

[0129] 각각의 샘플링 포인트에 있어서, 유의 수준($p < 0.05$)의 스튜던트의 t 검정을 이용하여, 각각, $A = [0, 53, 184, 1325, 1327, 738, 980, 1263, 915]$ 의 디퍼렌셜 유전자를 선출하였다.

[0130] 그리고, 선출한 디퍼렌셜 유전자의 셋트 A에 대하여, 오발현율(FDR) 및 2배 변경 스크리닝을 이용함으로써, 각 샘플링 포인트에 대하여, 각각, $B = [0, 29, 72, 195, 269, 163, 173, 188, 176]$ 의 유전자를 선출하였다.

[0131] 선출한 유전자 셋트 B에 대하여, 각 샘플링 포인트에 있어서, 상관이 높은 것을 하나의 클러스터로 통합하는 클러스터화를 행하여, 각각, 40개의 클러스터를 얻었다.

[0132] 또한, 각 샘플링 포인트에 있어서, 얻어진 40개의 클러스터 내의 모든 유전자에 대하여, 데이터의 정규화를 행한다. 그리고, 각 샘플링 포인트에 있어서, 정규화된 컨트롤 그룹 및 케이스 그룹에 있어서의 각 클러스터의 각각의 표준 편차의 평균값(SDd)(제3 지수), 클러스터의 멤버간의 피어슨의 상관 계수의 절대값의 평균값(PCCd)(제2 지수), 클러스터의 멤버와 다른 유전자 간의 피어슨 상관 계수의 절대값의 평균값(OPPCd)(제1 지수), 및

종합 지수(I)를 계산한다.

- [0133] 그리고, 각 샘플링 포인트에 있어서, 계산한 케이스 그룹에 있어서의 각 클러스터로부터, 종합 지수(I)가 가장 큰 클러스터를 DNB의 후보로서 선출하고, 또한, 상기 DNB의 후보에 대하여, 컨트롤 그룹의 표준 편차의 평균값(SDc), 유전자간의 피어슨 상관 계수의 절대값의 평균값(PCCc)을 기준으로, 유의성 분석에 의해, DNB인지의 여부를 판단한다. 그 결과, 각 샘플링 포인트에 있어서, DNB가 되는 클러스터의 수는, [0, 0, 0, 0, 1, 0, 0, 0, 0]이었다.
- [0134] 즉, 제5 샘플링 포인트(8 h)에서 DNB가 검출된, 상기 DNB는 220개의 유전자를 갖는 제111번째의 클러스터이다.
- [0135] 도 14a는 제1 검증예에 있어서, 검출된 DNB의 후보의 표준 편차의 평균값(SDd)의 시계열 변화의 예를 나타내는 그래프이다. 도 14b는 제1 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버간의 피어슨 상관 계수의 절대값의 평균값(PCCd)의 시계열 변화의 예를 나타내는 그래프이다. 도 14c는 제1 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버와 다른 유전자 간의 피어슨 상관 계수의 절대값의 평균값(OPCCd)의 시계열 변화의 예를 나타내는 그래프이다. 도 14d는 제1 검증예에 있어서, 검출된 DNB의 후보의 종합 지수(I)의 시계열 변화의 예를 나타내는 그래프이다. 도 14a~도 14d에 있어서, 횡축은 시간 단계(t)를 나타내고, 종축은, 각각, 표준 편차의 평균값(SDd)(도 14a), 클러스터의 멤버간의 피어슨 상관 계수의 절대값의 평균값(PCCd)(도 14b), 클러스터의 멤버와 다른 유전자 간의 피어슨 상관 계수의 절대값의 평균값(OPCCd)(도 14c), 및 종합 지수(I)를 나타내고 있다(도 14d). 파선은, 케이스 그룹으로부터 검출한 DNB의 후보의 각종 지수의 경시 변화를 나타내고, 실선은, 컨트롤 그룹에서 선택한 1 클러스터의 각종 지수의 경시 변화를 나타내고 있다.
- [0136] 도 14a~도 14d로부터 알 수 있듯이, 제4 시간 단계(즉, 4시간 경과)로부터, DNB의 후보의 제1 지수(PCCd), 제3 지수(SDd) 및 종합 지수(I)는 크게 증가하기 시작하여, 제5 시간 단계(즉, 8시간 경과)에서 피크에 달하고 있다. 한편, DNB의 후보의 제3 지수(OPCCd)는 제2 시간 단계로부터 저감하여, 마찬가지로 제5 시간 단계(즉, 8시간 경과)에서 극소값을 나타내고 있다.
- [0137] 또한, DNB의 동적 특성을 직관적으로 나타내기 위해, DNB를 포함하는 전체 유전자 네트워크의 동적 특성을 도 15에 나타낸다. 도 15는 제1 검증예에 있어서, 케이스 그룹의 유전자에 의해 구성된 네트워크에 있어서의 DNB의 동적 특성의 일례를 경시적으로 나타내는 맵이다. 도 15에는, 순서대로 0.5, 1, 4, 8, 12, 24, 48, 72 h의 각 샘플링 포인트에서, 케이스 그룹의 유전자 네트워크(3452개의 유전자, 9238개의 링크)가 표시되어 있고, 「○」로 표시하고 있는 노드는 DNB의 후보에 속하는 유전자이며, 「□」로 표시하고 있는 노드는 DNB의 후보의 노드 부근의 다른 유전자이고, 노드와 노드 간의 라인은 양 노드의 상관성을 나타내고 있다. 또한, 「○」의 색의 농도가 유전자의 표준 편차(SD)의 크기를 나타내고, 양 노드의 접속 라인의 농도가 양 노드의 상관 계수(PCC)의 절대값의 크기를 나타내고 있다. 도 15에 예시한 모든 맵은, 데이터 해석에 관한 오픈 소스 플랫폼인 Cytoscape(<http://www.cytoscape.org/>)를 이용하여 구성된 것이다.
- [0138] 도 15에 나타내는 바와 같이, DNB의 후보의 특성(SD, PCC)이 시간의 경과와 함께 변화하여, 다른 유전자와 같은 거동을 하는 통상의 클러스터로부터, 점차로 DNB로 진화해 간다. 도 15의 e에 나타내는 제5 단계(8 h가 경과한 시간)에서, DNB로서의 특징이 가장 현저하고, 질병 전 상태(8 h)인 경고 신호를 분명히 나타내고 있다. 그러나, 질병 상태로 천이한 후(24 h, 48 h, 72 h), DNB의 멤버는, 제차 다른 유전자와 동일한 거동을 하도록 되어 있다.
- [0139] 상기 결과로부터, 질병 전 상태는 제5 시간 단계의 근처에 존재하고, 제5 시간 단계 후, 시스템은 질병 상태로 천이한다고 하는 것이 판명된다.
- [0140] 따라서, 본 발명의 DNB에 의한 질병의 조기 진단 방법으로 진단하는 경우, 제4 시간 단계에서의 진단에서, 질병이 되는 경고 신호가 조금 보이고 있기 때문에, 머잖아 질병이 악화한다고 하는 진단 결과를 낼 수 있다. 그리고, 제5 시간 단계에서의 진단에서는, 질병이 되는 경고 신호를 분명히 보이고 있기 때문에, 곧 질병이 된다는 진단 결과를 낼 수 있다.
- [0141] 한편, 실제의 마우스 실험의 결과에서는, 케이스 그룹의 마우스는, 포스겐을 흡입하고 나서 8시간 후에, 폐수종을 일으켜, 12시간 후에 50%~60%가 사망하고, 24시간 후에, 더욱 60%~70%가 사망하였다.
- [0142] 따라서, 본 발명의 DNB에 의한 조기 진단의 진단 결과는, 마우스의 실제의 병 악화 상황과 완전히 일치하고 있다고 할 수 있다.
- [0143] <제2 검증예>

- [0144] 전술한 제1 검증예에 있어서, 동물 실험의 데이터를 이용하여, 본 발명의 DNB에 의한 질병의 조기 진단 방법의 유효성을 검증하였다. 본 검증예에서는, B 세포 림프종의 임상 데이터를 이용하여, 더욱, 본 발명의 DNB에 의한 질병의 조기 진단 방법의 진단 정밀도를 검증한다.
- [0145] 도 16은 제2 검증예에 있어서의 진단 데이터를 리스트한 테이블이다. 도 16에 나타내는 바와 같이, 임상상, 병변 및 플로우 사이토메트리에 기초하여, 샘플은, 안정기(P1), 활성화기(P2), 한계기(P3), 전이기(P4), 침공기(P5)의 5개의 단계의 그룹으로 나누어져 있다. 각 단계에 있어서, 샘플의 수는, 각각, 5, 3, 6, 5, 7이다. 각 단계에 있어서, 비종은, 각각, 「None」, 「None」, 「+/-」, 「+」, 「+++」이다. 각 단계에 있어서, 플로우 사이토메트리는, 각각, 「정상 안정」, 「정상 활성화」, 「이상」, 「믹스드」, 「B-1 클론」이다. 또한, 안정기(P1)에서 채취한 샘플을 컨트롤 샘플로 하고, 다른 단계(P2~P5)에서 채취한 샘플을 케이스 샘플로 한다.
- [0146] 상기 26개의 샘플로부터 측정된 유전자 발현 데이터로부터 13712개의 유전자를 대상으로, 전술한 DNB에 의한 질병의 조기 진단 방법으로 진단을 행하였다. 그 결과로서, 도 17a~도 17d에, 케이스 그룹의 유전자로부터 검출된 DNB의 후보의 각 지수를 나타낸다. 도 17a는 제2 검증예에 있어서, 검출된 DNB의 후보의 표준 편차의 평균값(SDd)의 시계열 변화의 예를 나타내는 그래프이다. 도 17b는 제2 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버간의 피어슨 상관 계수의 절대값의 평균값(PCCd)의 시계열 변화의 예를 나타내는 그래프이다. 도 17c는 제2 검증예에 있어서, 검출된 DNB의 후보의 클러스터의 멤버와 다른 유전자 간의 피어슨 상관 계수의 절대값의 평균값(OPCCd)의 시계열 변화의 예를 나타내는 그래프이다. 도 17d는 제2 검증예에 있어서, 검출된 DNB의 후보의 종합 지수(I)의 시계열 변화의 예를 나타내는 그래프이다.
- [0147] 도 17a~도 17d에 있어서, 횡축은 단계(P1~P4)의 번호를 나타내고 있으며, 종축은, 각각, 표준 편차의 평균값(SDd)(도 17a), 클러스터의 멤버간의 피어슨 상관 계수의 절대값의 평균값(PCCd)(도 17b), 클러스터의 멤버와 다른 유전자 간의 피어슨 상관 계수의 절대값의 평균값(OPCCd)(도 17c), 및 종합 지수(I)를 나타내고 있다(도 17d).
- [0148] 도 17a~도 17d에서 분명한 바와 같이, 제2 단계(P2), 즉 활성화기에 있어서, DNB의 후보의 종합 지수(I)가 피크에 달하고 있으며, 질병 상태를 나타내는 경고 신호가 가장 강하게 되어 있다. 이 진단 결과는, 실제의 병변과 완전히 일치하고 있다. 실제의 임상 데이터라도, 활성화기의 직후의 한계기로부터 병상이 악화하여, 비종이 「+/-」가 되고, 플로우 사이토메트리가 「이상」으로 되어 있다. 따라서, 본 검증예에 있어서의 DNB 특정 해석 결과와 실제의 임상 데이터와 완전히 일치하고 있다.
- [0149] 또한, 통상의 진단에서는, 도 16에 나타내는 바와 같이, 활성화기에 있어서의 비종이 「None」, 플로우 사이토메트리가 「정상 활성화」으로 되어 있기 때문에, 이상이 확인되지 않는다고 하는 진단 결과가 된다. 그에 대하여, 본 발명의 DNB에 의한 질병의 조기 진단 방법으로 진단하는 경우, 활성화기에 있어서, 질병 전 상태를 나타내는 경고 신호(DNB)를 검출하였기 때문에, 「이상의 징후가 보인다」라고 하는 진단 결과를 환자에게 전할 수 있다. 그 때문에, 환자는 조기 단계에서 치료 조치를 취함으로써, 병상의 악화를 멈춘다.
- [0150] 따라서, 본 발명의 DNB에 의한 질병의 조기 진단은, 림프종과 같은 복잡 질병의 조기 진단에 대단히 유효한 것이 검증되어 있다.
- [0151] 또한, 본 검증예에 있어서 검출된 DNB 중에, 22개의 유전자와 TFs가 포함되어 있고, 그 중, 13개의 유전자가 분명히 B 세포 림프종과 관련하고 있으며, 또한, 상기 13개의 유전자 중 8개가 증식의 마스터 레귤레이터인 것이 판명되어 있다. 따라서, 본 발명에 있어서의 DNB는, 질병 전 상태를 나타내는 경고 신호로서, 조기 단계에서 이상의 징후를 환자에게 전할 수 있을 뿐만 아니라, 질병과 관련하고 있는 유전자도 구체적으로 특정할 수 있기 때문에, 복잡 질병의 치료, 제약에 매우 도움이 된다고 생각된다.
- [0152] 상기 실시형태는, 본 발명의 무수하게 존재하는 실시예의 일부를 개시한 것에 지나지 않고, 질병의 종류, 검출하여야 하는 목적 등의 여러가지 요인을 가미하여 적절하게 설계 변경하는 것이 가능하다. 특히, 인자 항목으로서, 생체에 관한 측정에 의해 얻어진 정보이면, 여러가지 측정 데이터를 이용할 수 있다. 예컨대, 전술한 유전자, 단백질, 대사물에 관한 측정 데이터에 한정되지 않고, CT 스캔 등의 측정 장치에서 출력되는 체 내 화상에 기초하여, 각 부위의 여러가지 상황을 수치화함으로써, 측정 데이터로서 이용하는 것이 가능하다. 또한, 화상 이외에도, 소리 또는 체 내로부터 발생하는 소리를 측정하여, 수치화한 뒤에 측정 데이터로서 이용하는 것도 가능하다.

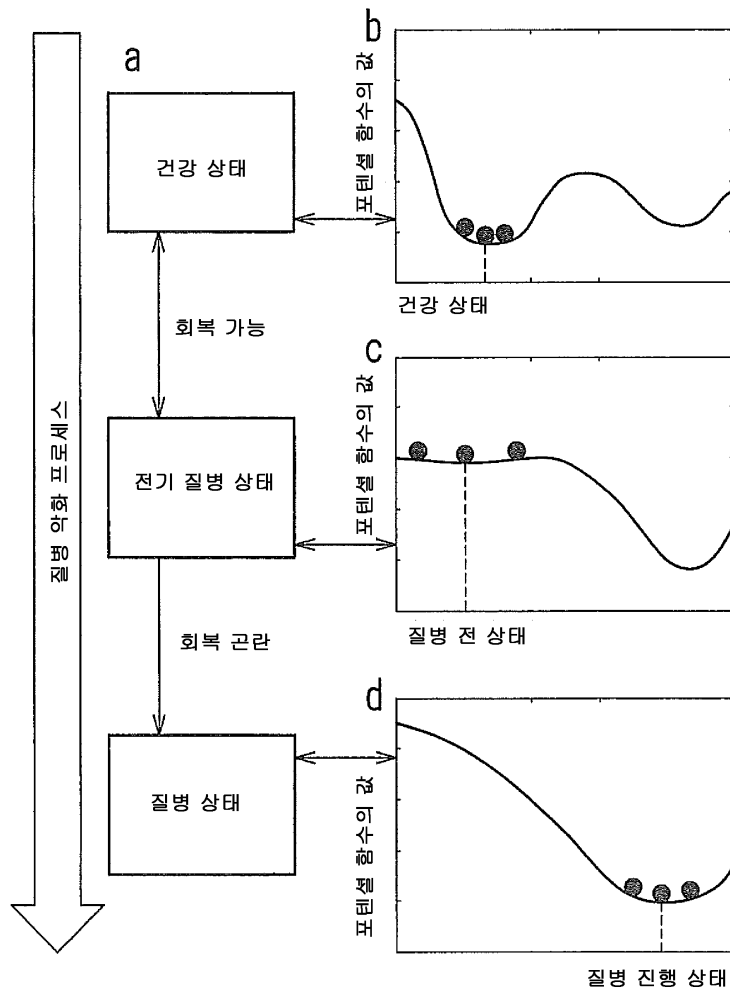
부호의 설명

[0153]

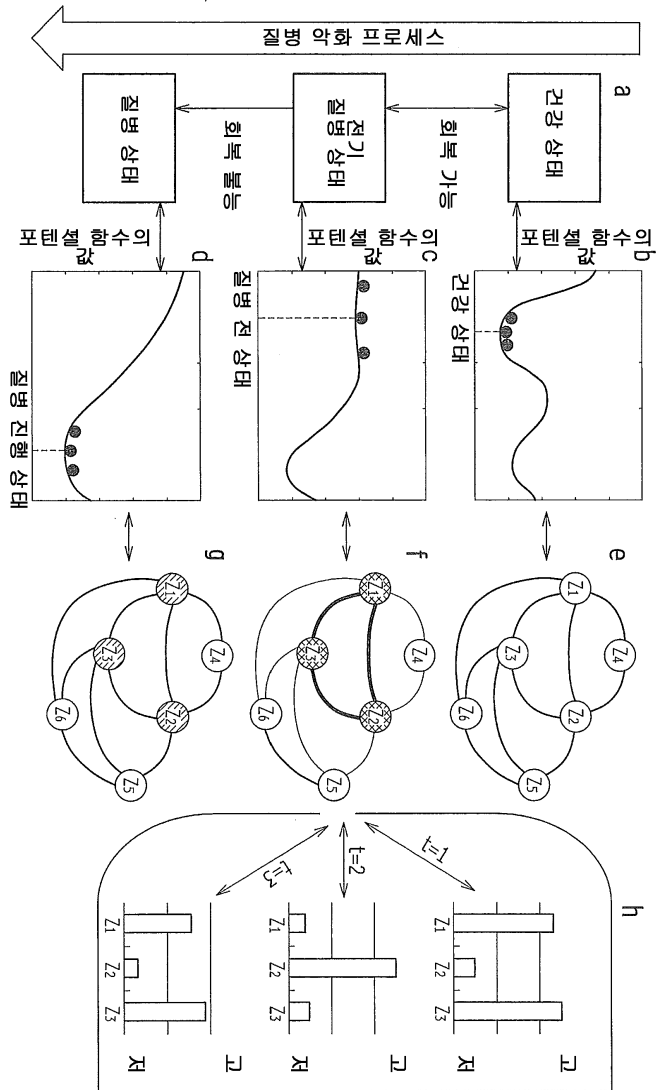
- 1 검출 장치
- 10 제어부
- 11 기록부
- 12 기억부
- 13 입력부
- 14 출력부
- 15 취득부
- 11a 검출 프로그램

도면

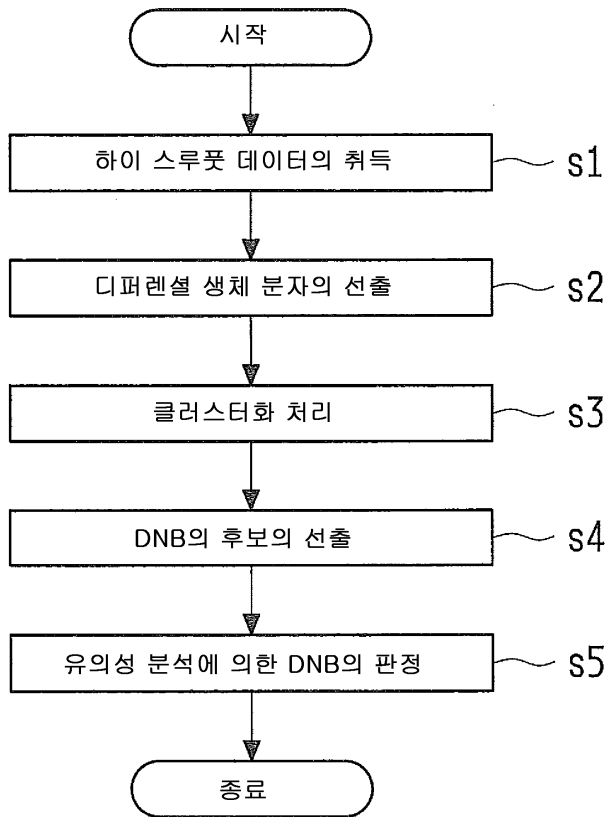
도면1



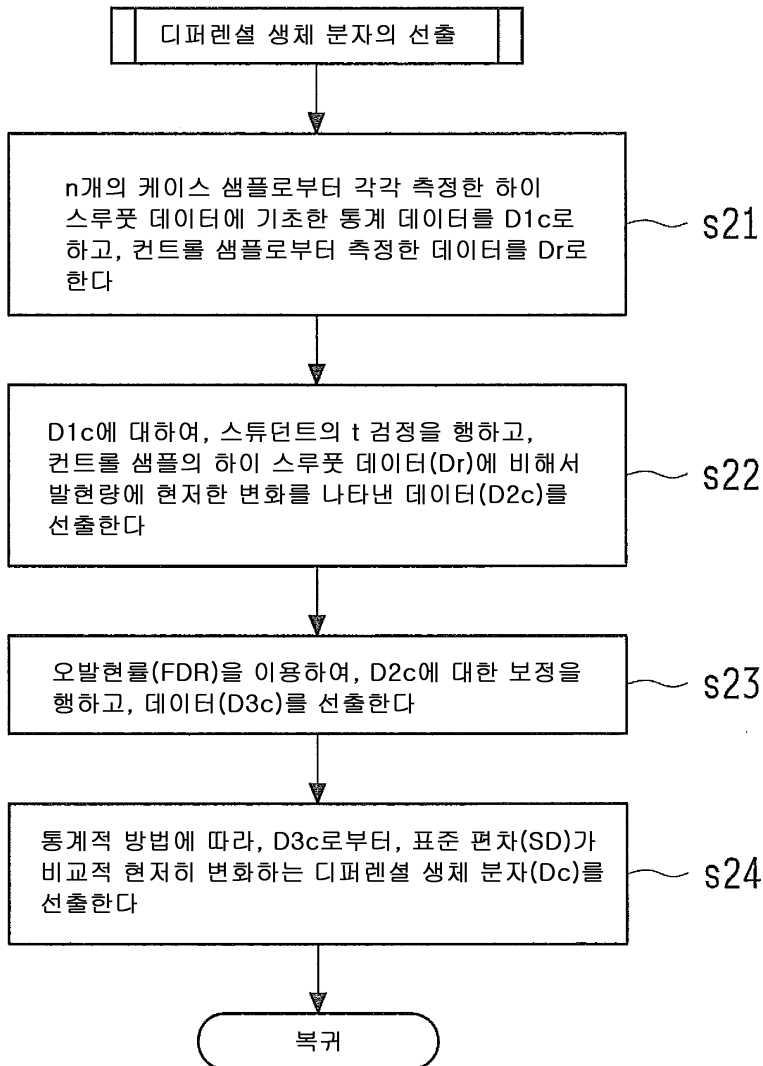
도면2



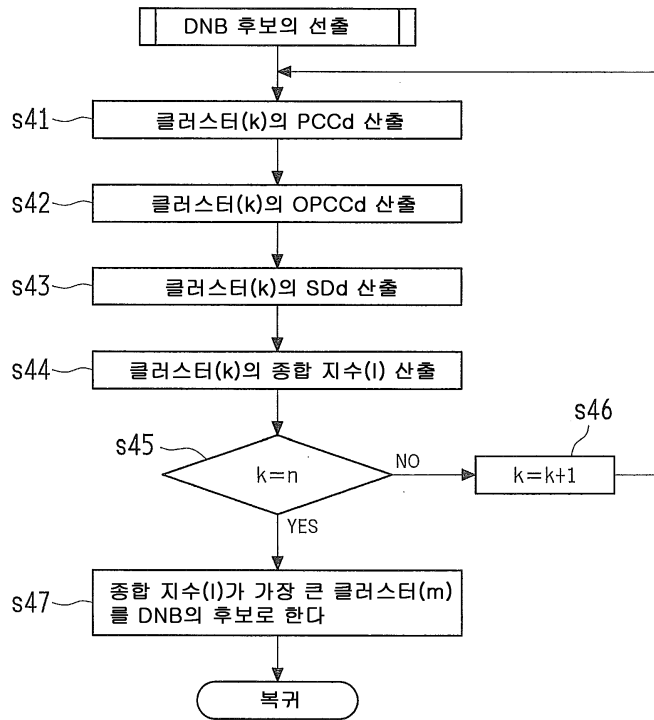
도면3



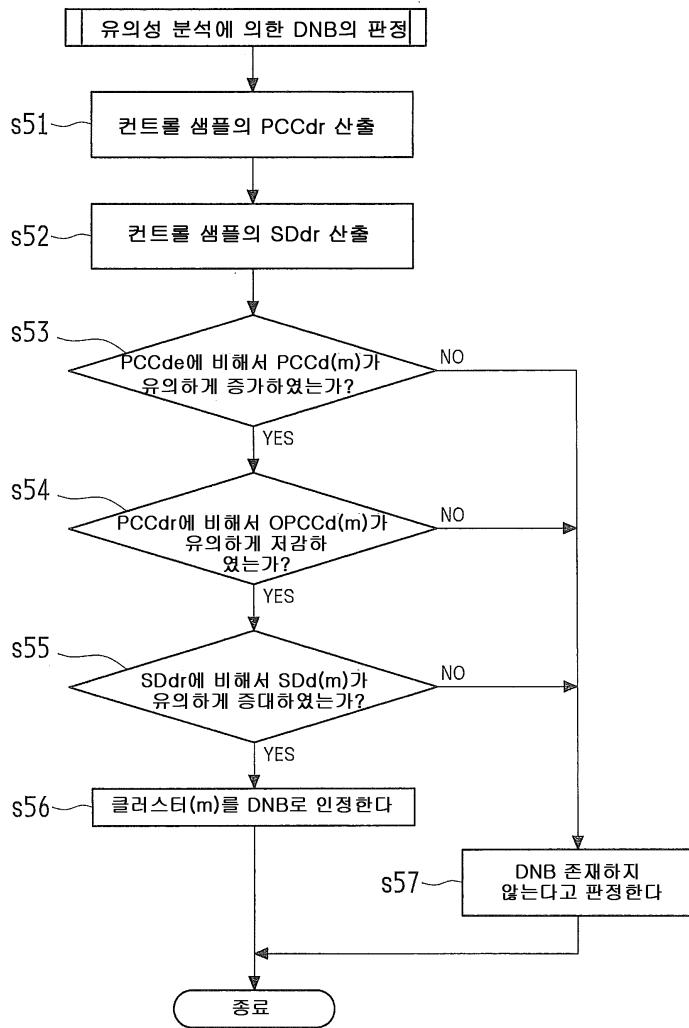
도면4



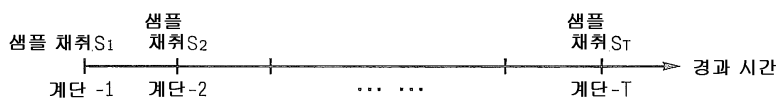
도면5



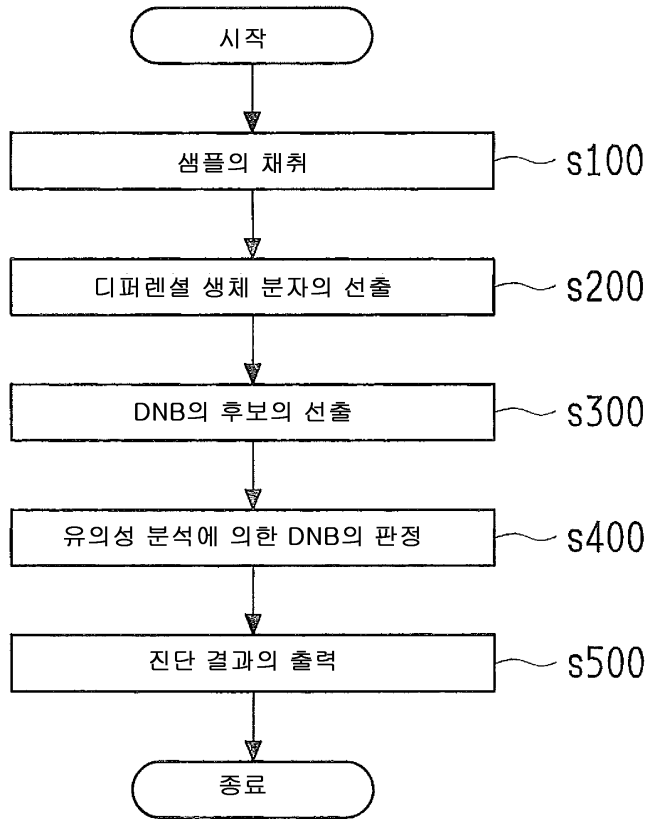
도면6



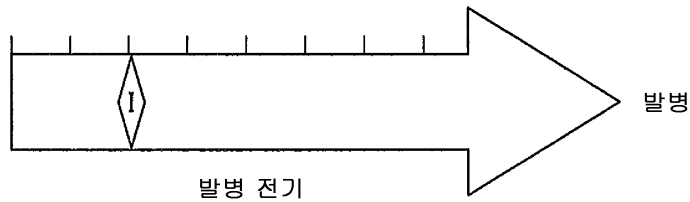
도면7



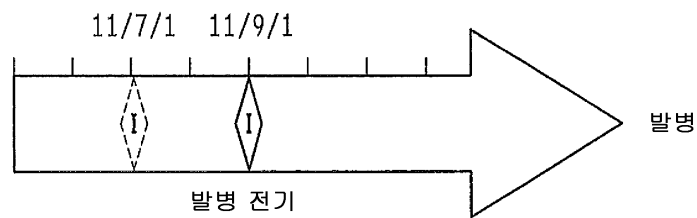
도면8



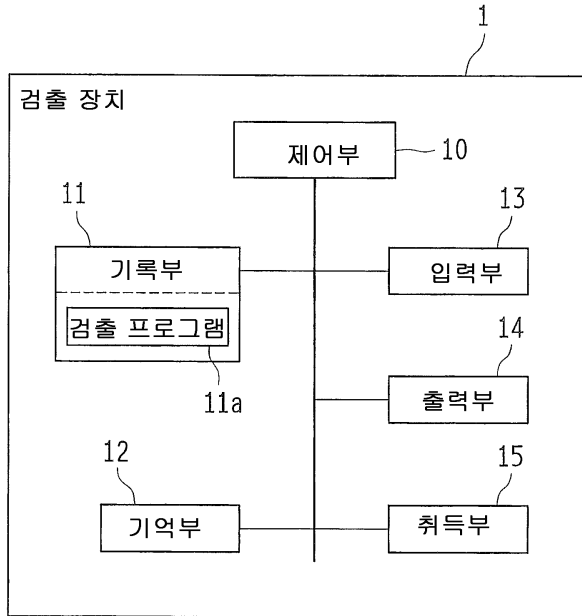
도면9



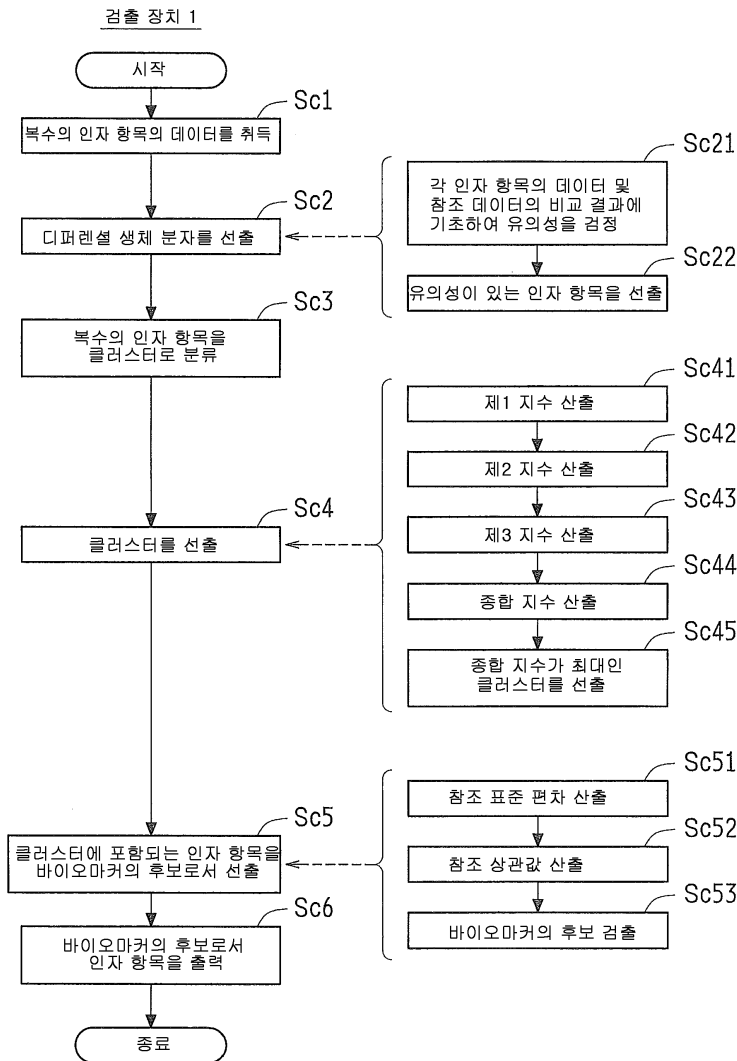
도면10



도면11



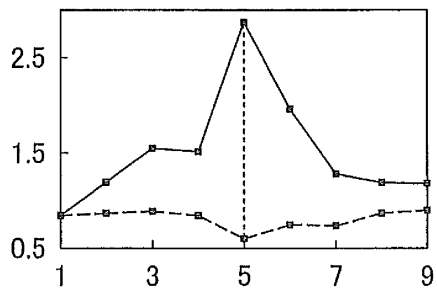
도면12



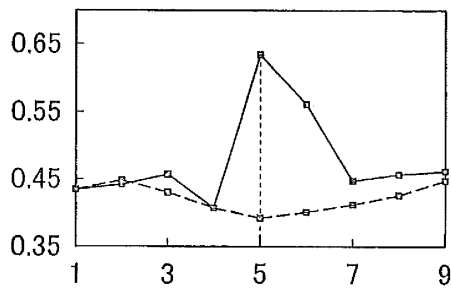
도면13

진단 대상	포스겐 유발성 폐 장애에 걸린 마우스(CD-1 수컷 마우스)
샘플 채집 대상	진단 대상으로 하는 케이스 그룹의 마우스의 폐 조직 참조 대상으로 하는 컨트롤 그룹의 마우스의 폐 조직
샘플링 포인트	0, 0.5, 1, 4, 8, 12, 24, 48, 72(h)
관찰 대상	22690개의 유전자

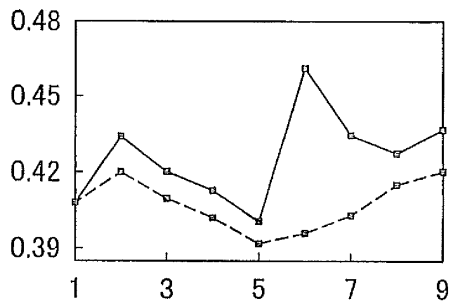
도면14a



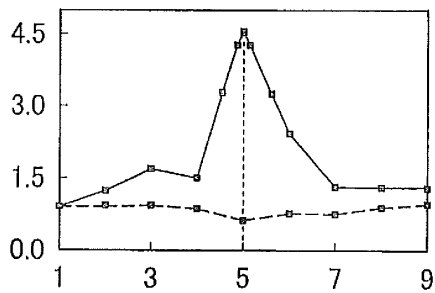
도면14b



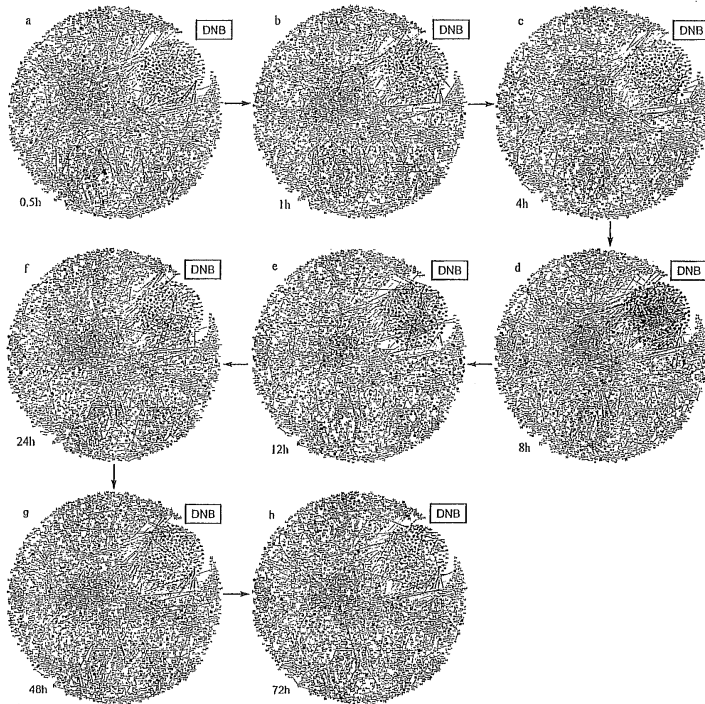
도면14c



도면14d



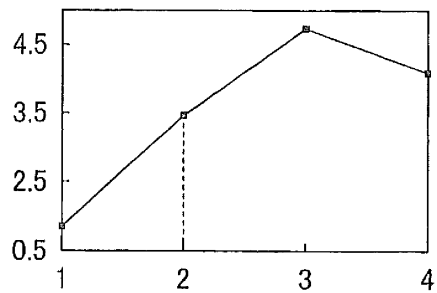
도면15



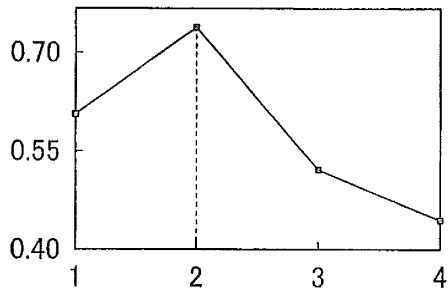
도면16

샘플 채집 대상	B 세포 림프종 환자			
	병변	샘플수	비중	플로우 사이토메트리
P1	안정기	5	None	정상 안정
P2	활성기	3	None	정상 활성
P3	한계기	6	+/-	이상
P4	전이기	5	+	믹스드
P5	침공기	7	+++	B-1 클론
케이스 그룹	P2~P5에서 채집한 샘플			
컨트롤 그룹	P1에서 채집한 샘플			
관찰 대상	13712개의 유전자			

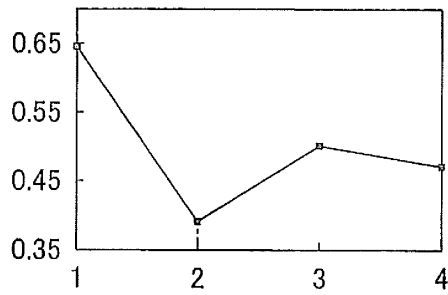
도면17a



도면17b



도면17c



도면17d

