



등록특허 10-2422494



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년07월19일
(11) 등록번호 10-2422494
(24) 등록일자 2022년07월14일

- (51) 국제특허분류(Int. Cl.)
G01N 33/68 (2006.01) *C12N 15/10* (2017.01)
- (52) CPC특허분류
G01N 33/6845 (2013.01)
C12N 15/1037 (2013.01)
- (21) 출원번호 10-2017-7017162
- (22) 출원일자(국제) 2015년11월24일
심사청구일자 2020년11월02일
- (85) 번역문제출일자 2017년06월22일
- (65) 공개번호 10-2017-0086627
- (43) 공개일자 2017년07월26일
- (86) 국제출원번호 PCT/GB2015/053573
- (87) 국제공개번호 WO 2016/083793
국제공개일자 2016년06월02일
- (30) 우선권주장
1420852.4 2014년11월24일 영국(GB)
- (56) 선행기술조사문헌
JP2012532607 A*
JP2013540440 A*
W02010135558 A1

*는 심사관에 의하여 인용된 문헌

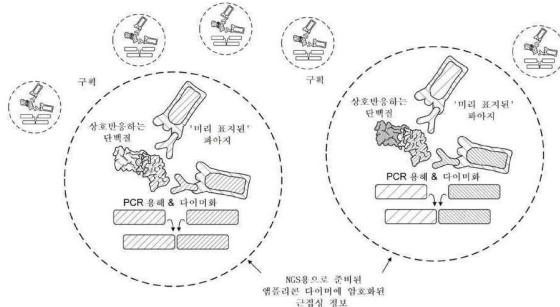
전체 청구항 수 : 총 31 항

심사관 : 차명훈

(54) 발명의 명칭 방법

(57) 요약

본 발명은 일반적으로 결합 상호작용, 특히 단백질-단백질 상호반응을 검출하기 위한 방법 및 키트, 상세하게는, 단백질-단백질 상호반응을 표지, 분석, 검출 및 측정하기 위한 고효율(high throughput) 방법에 관한 것이다.

대 표 도

(52) CPC특허분류

G01N 2458/10 (2013.01)

G01N 2500/04 (2013.01)

명세서

청구범위

청구항 1

결합체와 표적 사이의 결합 상호반응을 측정하는 방법으로,

a) 결합체 라이브러리의 각 멤버(member)가 고유의 뉴클레오티드 서열과 연관되어 있는 결합체 라이브러리와 표적을 접촉시켜 결합체/표적 복합체가 형성되도록 하는 단계로,

상기 결합체 라이브러리의 각 멤버와 연관되어 있는 상기 고유의 뉴클레오티드 서열은 상기 결합체 라이브러리의 각 멤버의 결합 특성과 연관되고,

상기 표적은 고유의 뉴클레오티드 서열과 연관되지 않는 단계;

b) 상기 복합체를 구획으로 단리(isolating)시켜 하나의 구획 내의 단일 복합체가 있도록하는 단계;

c) 상기 각 복합체에서 상기 결합체(들)와 연관되어 있는 고유의 뉴클레오티드 서열을 연결(linking)하여 연결된 뉴클레오티드 서열을 형성시키고 상기 단리된 복합체의 단리를 이 연결 단계 중에 유지시키는 단계;

d) 상기 연결된 뉴클레오티드 서열로부터 각 복합체에 존재하는 결합체(들)를 식별하는 단계를 포함하는 방법.

청구항 2

결합체와 표적 사이의 결합 상호반응을 측정하는 방법으로,

a) 결합체 라이브러리를 표적과 접촉시켜 결합체/표적 복합체를 형성시키고, 여기서, 상기 결합체 라이브러리의 각 멤버(member)는 고유의 뉴클레오티드 서열과 연관되고, 상기 표적이 고유의 뉴클레오티드 서열과 연관되도록 하는 단계;

b) 상기 복합체를 구획으로 단리(isolating)시켜 하나의 구획 내에 단일 복합체가 있도록하는 단계;

c) 상기 결합체 및 각 복합체 내의 표적과 연관되어 있는 고유의 뉴클레오티드 서열(들)을 연결(linking)하여 연결된 뉴클레오티드 서열을 형성시키고 상기 단리된 복합체의 단리를 이 연결 단계 중에 유지시키는 단계;

d) 상기 연결된 뉴클레오티드 서열로부터 각 복합체에 존재하는 결합체(들)를 식별하는 단계; 및

e) 상기 결합체 라이브러리의 각 멤버(member)의 고유한 뉴클레오티드 서열을 상기 멤버의 결합 특성과 연관시키는 연결된 뉴클레오티드를 사용하는 단계;를 포함하는 방법.

청구항 3

제1항 또는 제2항에 있어서, 상기 결합체 라이브러리가 항체 라이브러리를 포함하는 방법.

청구항 4

제3항에 있어서, 상기 항체 라이브러리가 항체 디스플레이 라이브러리 또는 항체의 라이브러리를 포함하는 방법으로, 각각의 항체는 상기 고유의 뉴클레오티드 서열로 표지된 방법.

청구항 5

제1항 또는 제2항에 있어서, 상기 표적이 단백질을 포함하는 것인 방법.

청구항 6

제1항 또는 제2항에 있어서, 상기 표적이 단백질 디스플레이 라이브러리를 포함하며, 여기서 상기 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있는 것인 방법.

청구항 7

제6항에 있어서, 상기 단백질 디스플레이 라이브러리가 cDNA 파아지 디스플레이 라이브러리인 방법.

청구항 8

제5항에 있어서, 상기 단백질이 단백질 혼합물 내에 있는 방법.

청구항 9

제8항에 있어서, 상기 단백질 혼합물이 강화된(enriched) 단백질 혼합물인 방법.

청구항 10

제8항에 있어서, 상기 단백질 혼합물이 인단백질(phosphoprotein), 막단백질(membrane protein), 및/ 또는 천연 또는 인공적으로 개량된 단백질인 방법.

청구항 11

제1항 또는 제2항에 있어서, 다른 결합제에 결합하는 결합제를 사용 전에 상기 결합제 라이브러리로부터 제거하는 방법.

청구항 12

제1항 또는 제2항에 있어서, 결합제 라이브러리를 사용 전에 강화시키는 방법.

청구항 13

제1항 또는 제2항에 있어서, 상기 구획이 고체 지지체 상의 세트 위치인 방법.

청구항 14

제1항 또는 제2항에 있어서, 상기 구획이 에멀젼 액적(emulsion droplet), 확산 제한되거나 분리된 구획(compartment)인 방법.

청구항 15

제1항 또는 제2항에 있어서, 상기 복합체 중의 결합제와 연관되어 있는 뉴클레오티드 서열이 상기 복합체 중의 결합제와 연관되어 있는 다른 뉴클레오티드 서열에 결합되는 방법.

청구항 16

제4항에 있어서, 상기 복합체 중의 결합제와 연관되어 있는 뉴클레오티드 서열이 결합제/표적 복합체 내의 표적과 연관되어 있는 뉴클레오티드 서열에 결합되는 방법.

청구항 17

제1항 또는 제2항에 있어서, 상기 복합체 내에 존재하는 결합제 및/또는 표적이 상기 연결된 뉴클레오티드 서열로부터 식별되는 방법.

청구항 18

제1항 또는 제2항에 있어서, 상기 복합체 내에 존재하는 (i) 결합제 또는 (ii) 결합제와 표적이 상기 연결된 뉴클레오티드 서열의 서열분석에 의해 식별되는 방법.

청구항 19

제1항 또는 제2항에 있어서, 상기 결합제/표적 복합체가 고정화된 다음 단리되는 방법.

청구항 20

제1항 또는 제2항에 있어서, 1개 이상의 복합체로부터 연결된 뉴클레오티드 서열들을 존재하는 결합제 및/또는 표적의 식별 전에 합하는 방법.

청구항 21

제1항 또는 제2항에 있어서, 결합체 라이브러리와 표적을 접촉시키는 단계 (a)를 화합물의 존재 또는 부재하에서 수행하는 방법.

청구항 22

제5항에 있어서, 상기 단백질이 단백질 혼합물 내의 다른 단백질과 가교-결합되는 방법.

청구항 23

제1항 또는 제2항에 있어서, 상기 단리를 희석, 특이적 결합, 또는 복합체의 물리적 및/또는 화학적 특성에 의한 분리로 수행하는 방법.

청구항 24

제3항에 있어서, 상기 항체 라이브러리는 항체 파아지 라이브러리이고, 상기 항체 파아지 라이브러리의 각 멤버와 연관되어 있는 고유의 뉴클레오티드 서열이 상기 항체의 CDR을 코딩하는 파아지 내의 뉴클레오티드 서열인 방법.

청구항 25

제7항에 있어서, 상기 cDNA 파아지 디스플레이 라이브러리의 각 멤버와 연관되어 있는 고유의 뉴클레오티드 서열이 디스플레이된 단백질을 코딩하는 파아지 내의 서열인 방법.

청구항 26

제1항 또는 제2항에 있어서, 상기 뉴클레오티드 서열을 결합하는 단계가:

- i. 결합체 또는, 결합체 및 표적과 연관되어 있는 뉴클레오티드 서열을 적어도 2쌍의 PCR 프라이머를 사용하여 증폭시켜 적어도 2세트의 앰플리콘을 생산하는 단계로, 상기 프라이머는 제1 세트의 프라이머에 의해 생성된 앰플리콘이 제2 세트의 프라이머에 의해 생성된 앰플리콘 중의 서열에 상보성인 서열을 포함하도록 디자인된 단계;
- ii. 앰플리콘 세트를 어니얼링(annealing)시키는 단계;
- iii. 증폭 반응을 수행하여 연결된 뉴클레오티드 서열을 생산하는 단계를 포함하는 방법.

청구항 27

제26항에 있어서, 상기 증폭이 애밀전 PCR 증폭법을 사용하여 수행되는 방법.

청구항 28

제26항에 있어서, 단계 i-iii이 동시에 수행되는 방법.

청구항 29

제1항 또는 제2항에 있어서, 상기 방법을 반복하고 표적 또는 결합체 라이브러리의 농도를 변화시키는 방법.

청구항 30

제1항 또는 제2항의 방법을 수행하기 위한 키트로,

- a) 결합체 라이브러리의 각 멤버(member)가 예정된 단백질 결합 특성을 가지고 알려진 고유의 뉴클레오티드 서열과 연관되어 있는 결합체 라이브러리; 및
- b) 상기 결합체와 연관되어 있는 뉴클레오티드 서열을 결합하기 위한, 적어도 2쌍의 프라이머 세트; 및 선택적으로 사용 설명서를 포함하는 키트.

청구항 31

제30항에 있어서, 라이브러리의 각 멤버가 고유의 뉴클레오티드 서열과 연관되어 있는 단백질 디스플레이 라이

브러리를 추가로 포함하는 키트.

청구항 32

삭제

청구항 33

삭제

발명의 설명

기술 분야

[0001]

본 발명은 일반적으로 결합 상호반응, 특히 단백질-단백질 상호반응을 검출하기 위한 방법 및 키트, 상세하게는, 단백질-단백질 상호반응을 표지, 분석, 검출 및 측정하기 위한 고효율(high throughput) 방법에 관한 것이다.

배경 기술

[0002]

세포의 아키텍처(architecture)는 이의 복합체인, 세포를 실제로 만드는 분자 기계(molecular machine)로 정의된다. 세포생물학은 전통적으로 촉매, 시그널링 분자, 또는 세포 및 미생물의 빌딩 블럭(building block)으로서 단백질 각각의 작용을 기반으로 이들을 식별한다. 현재, 우리는 단백질-단백질 상호반응의 네트워크에서의 요소(element)로서 이에 대해서뿐만 아니라, 기능적 모듈 내에서 '컨텍스트적(contextual)' 또는 '세포의(cellular)' 기능에 대하여, 단백질의 역할을 확장하는 포스트-게놈 관점의 출현을 목격하였다.

[0003]

복합 단백질-단백질 네트워크의 정성적 및 정량적 특성화 및 주요 세포 타입 특이적 상호반응 단백질의 식별은 암, 자가면역 질환 및 기타 질병과 같은 다수의 인간의 질병에서 단백질-단백질 상호반응의 생리학적 과정과 변화를 이해하는데 있어서 무엇보다도 중요하다. 단백질-단백질 네트워크에서의 상세한 고찰 및 질환-연관 차이의 식별이 특이적 약물의 합리적인 디자인 및 개발을 위한 새로운 방법으로 이어질 수 있다. 세포 또는 조직에서 단백질-단백질 상호반응의 패턴이 또한 분자 진단용 도구(tool)로 사용될 수 있다.

[0004]

단백질은 세포의 수많은 생리학과 기능에 대한 역학적 토대(mechanistic foundation)를 대표하는 복합 상호반응에 참여한다. 이를 단백질-단백질 상호반응은 복합 네트워크로 정교하게 정리된다. 단백질-단백질 상호반응 네트워크의 아키텍처는 척도가 없는(scale-free) 것으로 제안되었으며, 대부분의 단백질은 연결부위(connection)를 1개 또는 2개만 갖지만 상대적으로 수가 더 적은 '허브(hub)'는 수십, 수백개 이상의 링크(link)를 갖는다. 상호반응 네트워크는 고도로 역동적이어서, 상호반응체(interactome)에서, 예를 들어 외부 자극체 또는 심지어 발달 과정에 대해 신속하게 변화가 일어나도록 한다. 코어 단백질(core protein) 간 및 2개 이상의 모듈 단백질 간의 상호반응은 도메인-도메인 상호반응에 의해 매개되는 것 같다. 부착 단백질 내 및 사이에서의 상호반응은 이런 방식으로 털 일어나는 것 같다. 단백질 복합체와 상호반응이 생물학적 과정의 조절 및 수행에 기여함에도 불구하고, 상대적으로 적은 수의 복합체들이 구조와 기능면에서 잘 알려져 있다.

[0005]

세포의 상호반응에 대한 운동상수를 실험적으로 얻기 위한 시도는 거의 없다. 이러한 정량적 파라미터는 세포 과정의 미분방정식-기반 운동 모델을 개발할 수 있도록 할 것이다. 그러한 모델은 약물 작용을 이해하는데 있어서 필수적이며 수많은 복합 질환에 대한 신규 약물의 발견을 촉진할 것이다. 정량적 다중-척도 모델의 개발은 세포 수준에서 약물의 치료 작용과 부작용의 이론적 이해를 제공할 수 있다.

[0006]

용어 '샘플링'은 서브세트(subset) 집단에서 정보를 얻는 실험 디자인에 대해 사용된다. 대표 샘플링(representative sampling)은 단백질 상호반응 데이터세트의 생성에서 통상적이지 않으며, 여기에서 샘플링은 흔히 생물학적 우선권에 의해 안내되었다. 가능한 상호반응의 총 세트 중 일부가 실제로 시험되어 '커버리지(coverage)'로 요약된다. 요즘 기술에 비추어, '상호반응체', 예를 들면, 연구될 조건하에서 세포에서 일어나는 모든 물리적 상호반응 세트에 대해 추론하는 것이 타당하지 않다.

[0007]

다른 단백질에 결합하는 단백질을 선택하고 검출하기 위한 물리적 방법을 포함하여 단백질-단백질 상호반응을 연구하기 위한 여러 가지 방법, 예로서 단백질 친화성 크로마토그래피, 친화성 블로팅법, 면역침전법(2D 젤 전기영동법 및 질량 분석법 포함), 가교-결합법; 라이브러리-기반(library-based) 방법: 단백질 프로팅(protein

probing), 파아지 디스플레이(phage display), 투-하이브리드(two-hybrid) 시스템, 다른 라이브러리-기반 방법 및 유전적 방법: 유전자 외의 억제인자(extragenic suppressor), 합성 치사효과, 과잉생산 표현형, 야생형 단백질의 과잉생산 및 돌연변이 단백질의 과잉생산; 및 연결되지 않은 비-상보성이 고안되었다.

[0008] 이들 방법 중 많은 것들이 고효율 단백질-상호반응 분석에 적합하지 않다. 가장 전망 있는 고효율 기술은 펩티드- 및 단백질-라이브러리 선별 기법, 예로서 관심의 대상인 단백질과 상호반응하는 단백질에 대한 유전자를 식별하고 복제하는 방법인 이스트 투-하이브리드(yeast two-hybrid) 전략법; 각각의 효모 콜로니가 리포터 유전자 활성 - 자동화된 방식으로 상호반응이 디스플레이됨 - 에 대해 점수를 매길 수 있는, 정의된 쌍의 '베이트(bait)'와 '먹이(prey)' 단백질을 발현하는, 콜로니-어레이 포매트에서 대-규모 실험이 수행되는, 투-하이브리드(yeast two-hybrid) 어레이; 특히 세포 ('콤플렉솜(complexome)') 내 모든 복합체 및 이들의 구성 단백질을 규정하기 위한, "베이트(bait)" 단백질에 대해 단백질의 라이브러리가 패닝(panning)되는 파아지 디스플레이 및 친화성-정제법/질량분석법(affinity-purification/mass-spectrometry, AP-MS); 및 텐덤 친화성 정제법(tandem affinity purification, TAP)으로 입수할 수 있다. TAP는 상호반응하는 단백질을 다양한 형태의 해당 복합체에서 이들의 출현 빈도에 따라서, 코어, 모듈, 또는 부착 단백질로 밝혀진다.

[0009] 이들 방법들은 모두 이들 기법을 사용함으로써 얻게 되는 정보의 신뢰도, 완성도 및 용이성과 관련하여 장점과 단점을 갖는다. 이상적인 방법은 시간 및 비용 효과적인 방식으로 상호반응체의 정보를 포착하여, 랜덤 샘플링 및 높은 리던던시(redundancy)의 샘플링이 가능하도록 한다. 이는 심지어 큰, 다중-유니트 단백질 복합체의, 동적이고, 원래의 세포 컨텍스트 기반의, 천연 단백질-단백질 상호반응 기반의, 포괄적이고 충분히 큰, 정량적 상호반응 데이터의 커버리지를 제공한다. 이는 비-특이적이고, 돌발적으로 상호반응하는 단백질의 검출과 같은, 랜덤 변수의 효과를 억제한다. 이는 또한, 원래의 단백질-단백질 상호반응을 제외한 검출 원리에 관여하는, 임의의 결합 반응 연관 변수인, 변수들의 효과를 없앤다.

[0010] 투-하이브리드(two-hybrid) 선별법, 특히 어레이(array) 기반 기법은 대규모 상호반응체 정보 생성을 가능케 한다. 그러나, 이들의 이중성(binary), 쌍을 이루는 방식의 검출법, 원래의 컨텍스트 기반 동적 정보의 부족, 인공 결합체(하이브리드 단백질) 및 효모 세포 컨텍스트 제한된 원리 (예를 들면, 원래의 숙주와 비교하여 왜곡된 번역후 변형)로 인한 주요 단점이 있다. 이들 방법 중 거의 모두가 다양한 방식에 의해 부분적으로 해결되었다. 그러나, 이들 요구되는 특성을 모두 합한 방법은 고안되지 않았다.

[0011] 친화성 기반 방법, 특히 검출 원리로서 질량분석법을 사용하는 것들은 높은 양의 반(semi)-정량적 상호반응체 데이터를, 부분적으로 정확한 세포의 컨텍스트로 생성시킨다. 그러나, 이들은 랜덤하고 결합하는 (친화성) 연관 변수에 의해 영향을 받는다. 이들은 돌발적이고, 비-특이적인 결합 반응을 검출한다. 랜덤하게 샘플링된, 높은 커버리지, 포괄적인 데이터세트(dataset)를 생성시키기 위해서는 현저한 양의 시간과 비용을 필요로 하며, 이는 상호반응체의 동적인 특성을 검출할 수 있는 이의 잠재적 이점을 훼손한다. 이러한 이슈 중 일부가 해결되었는데, 특히 텐덤 친화성 정제법(TAP)을 사용함으로써 해결되었으며, 이 방법은 돌발적이고, 비-특이적인 결합 반응을 최소로 감소시키지만, 신뢰도가 약한 단백질-단백질 복합체 회수를 훼손시킨다.

[0012] 이러한 기법들은 단백질-단백질 상호반응(protein-protein interaction, PPI) 데이터의 생성을 대규모로 가속시켰다. 상호반응체에 대한 개척 연구 후, 몇몇 대규모 연구가 수행되어 쌍을 이루는 방식의 단백질-단백질 상호반응의 고품질 데이터세트가 일부 생성되었다. 예를 들면, 여과된 효모 상호반응체(filtered yeast interactome, FYI)는 Y2H 데이터, AP-MS 데이터, *in-silico* 예측, Munich Information Centre for Protein Sequences 물리적 상호반응, 및 문헌에 보고된 단백질 복합체를 포함한, 상이한 데이터세트의 교집합이다.

발명의 내용

해결하려는 과제

[0013] 현존하는 방법론적 접근법은 단백질-단백질 상호반응 및 상호반응체(interactome) 연구의 니즈(needs)를 완전하게 만족시키지 못하기 때문에, 복합 단백질-단백질 네트워크의 분석 및 특성화를 위한 새로운 방법이 요구된다.

과제의 해결 수단

[0014] 본 발명은 결합 상호반응, 특히 단백질-단백질 상호반응을 세포 수준(cellular level)으로 검출하기 위한 방법 및 키트를 제공한다. 상기 방법 및 키트는 복합 단백질 네트워크, 바람직하게는 세포의 원래의 컨텍스트(original context)에서 모든, 또는 서브세트의 상호반응하는 단백질을 동시에 검출하는데 사용될 수 있다. 상

기 방법 및 키트는 심지어 큰, 다중-유니트 단백질 복합체의 정량적이며 잠재적으로 운동에 의한 상호반응 데이터의 동적이며, 원래의 세포 컨텍스트(original cellular context) 기반의, 천연 단백질-단백질 상호반응 기반의, 포괄적이며, 충분히 큰 커버리지를 제공한다.

[0015] 본 발명은 결합제(binding agent)로서 복수의 항체 파아지(phage)를 사용하는, 항체 디스플레이 기술을 사용하여 단백질-단백질 상호반응을 검출하는데 사용될 수 있다. 본 발명은 또한, 결합제로서 복수의 압타머(aptamer)를 사용하는, 압타머 기술을 사용하여 단백질-단백질 상호반응을 검출하는데 사용될 수 있다. 복수의 결합제의 복잡성(complexity)은 몇개의 결합제와 수만 또는 수십만 또는 수백만 또는 수천만 또는 수억 개의 결합제 사이에서 광범위하게 변화될 수 있다. 본 발명된 방법에 적합한 복잡성이 높은 결합제로부터 복잡성이 낮은 결합제를 얻기 위하여, 복잡성 감소법이 고안되었다(강화(enrichment)).

[0016] 표적 분자 사이에서 더욱 상세한 상호반응을 식별하고 모니터할 수 있다. 예를 들어, 단백질-단백질 상호반응을 검출할 수 있다. 결합제/표적 복합체(agent/target complex) 내에 2개 이상의 결합제가 존재한다는 것은 복합체 내에 2개 이상의 표적이 존재할 수 있다는 것을 나타낼 수 있다. 이는 2개 이상의 표적이 서로 상호반응하거나 결합될 수 있음을 나타낸다. 특이적 결합제의 식별 가능한 부분, 예를 들어 단백질 또는 핵산 서열이 알려져 있는 경우, 표적을 식별할 수 있다. 이 방법은 결합되어 디스플레이된 항체 파아지의 식별 가능한 핵산 서열, 즉, 예정된 결합 특성을 갖는 것들, 예를 들면, 공지되어 있는 에피토프 서열을 갖는 것들, 또는 특이적 분자에 결합하는 것으로 공지되어 있는 것들을 연결시킴으로써 고도의 평행 PCR(parallel PCR) 증폭법을 사용하여 수행될 수 있다. 이는 바람직하게는 에멀션 PCR(emulsion PCR)에 의해 수행될 수 있다. 이는 낮은 단백질 복합체 농도로, 바람직하게는 구획(compartment)에서 수행될 수 있다. 표적 간의 상호반응, 예를 들어, 단백질-단백질 상호반응은 고도의 평행 PCR 증폭법에 의해, 바람직하게는 복잡성이 감소된 결합 검출제를 사용하여 검출될 수 있다. 표적-표적, 예를 들면, 단백질-단백질 상호반응 정보는 연결된 식별 가능한 서열의 서열분석에 의해, 바람직하게는 고도의 평행 DNA 서열분석에 의해 또는 기타 서열 검출 수단에 의해 얻어진다. 투입 물질, 예를 들면 표적의 양을 변화시킴으로써 리간드 결합 운동역학 데이터를 수집할 수 있다. 또한, 상기 방법은 화합물이 표적 상호반응에 대해 임의의 효과를 갖는지, 그리고 이런 효과가 작용성(agonistic)인지 길항성(antagonistic)인지 결정하기 위하여 상기 화합물의 존재 및 부재하에서 수행될 수 있다.

[0017] 본 발명은 또한 유기체의 단백질 단편(fragment)을 디스플레이하고 다수의 디스플레이된 항체의 결합 특성을 측정하는 단백질 디스플레이 기술을 사용할 수 있으며, 각각의 항체는 고유의 식별 가능한 서열 정보를 가지며 각각의 디스플레이된 단백질 단편은 식별 가능한 서열 정보를 갖는다. 바람직하게는, 상기 디스플레이된 단백질 단편의 식별 가능한 서열 정보가 디스플레이된 아미노산 서열을 코딩하는 서열이다. 결합된 항체의 동일성(identity)은 각각의 항체-단백질 복합체에 대해 식별 가능한 서열 정보로부터 결정될 수 있다. 선택적으로, 각각의 항체-단백질 복합체 내에서, 결합된 단백질 단편의 동일성이 식별될 수 있다. 선택적으로, 결합된 항체의 동일성과 결합된 단백질 단편의 동일성이 각각의 항체-단백질 복합체에 대한 연결된 식별 가능한 서열 정보로부터 결정될 수 있다. 상기 결합, 운동역학적 특성이 또한 상이한 양의 표적, 예를 들면, 디스플레이된 단백질 또는 디스플레이 항체와 같은 단백질 및 결합제를 사용하여 결정될 수 있다.

발명의 효과

[0018] 본 발명의 방법 및 조성물은 또한 그러한 단백질-단백질 상호반응을 작용시키거나 길항시킬 수 있는 화합물을 식별하는데 사용될 수 있다. 본 발명은 길항적(방해하는) 또는 작용적(촉진하는) 화합물과의 결합 상호반응을 검출하기 위한 방법 및 키트를 제공한다. 본 발명은 복합 단백질 네트워크에서, 바람직하게는 세포의 원래의 컨텍스트에서 길항적 및/또는 작용적 화합물의 결합 상호반응을 동시에 검출하기 위한 방법 및 키트를 제공한다. 상기 방법 및 키트는 원래의 세포 컨텍스트 기반의, 천연 단백질-단백질 상호반응 기반 데이터를 제공하고, 이는 포괄적이며, 정량적이고, 잠재적으로, 운동역학적 상호반응 데이터, 심지어 큰, 다중-유니트 단백질 복합체에 대해 충분히 큰 커버리지를 갖는다.

[0019] 본 발명은 결합제와 표적 사이의 결합 상호반응을 측정하는 방법을 제공하며, 이 방법은

[0020] a) 결합제 라이브러리와 표적을 접촉시켜 결합제/표적 복합체가 형성되도록하는 단계 (여기서 상기 결합제 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다);

[0021] b) 상기 결합제/표적 복합체를 분리시키는 단계;

[0022] c) 결합제/표적 복합체에서 결합제와 연관되어 있는 뉴클레오티드 서열을 연결하여 연결된 뉴클레오티드 서열을

형성시키는 단계;

[0023] d) 상기 연결된 뉴클레오티드 서열로부터 복합체에 존재하는 결합체를 식별하는 단계를 포함한다.

[0024] 본 발명은 복합체 결합 상호반응, 특히 단백질-단백질 네트워크 또는 상호반응체(interactome)를 분석하고 특성화하는 방법을 기술한다. 상기 방법은 결합체 및 선택적으로 이들의 표적, 예컨대 단백질의 동시-국소화(co-localization)에 기반하고 있으며, 여기서 결합체 및 선택적으로 이들의 표적의, 바람직하게는 복수 개의 구획에서의 동시-국소화에 대한 정보는 쌍을 이루는 방식으로 연결되고 뉴클레오티드로 번역된다. 상기 결합체 및 선택적으로 표적의 동일성이 또한 상기 뉴클레오티드 서열로부터 결정될 수 있다. 이 정보는 서열분석으로 밝혀질 수 있다.

[0025] 본 발명은 또한 표적 분자 상호반응에 대한 길항적(방해하는) 또는 작용적(촉진하는) 화합물의 효과를 분석하고 특성화하는 방법을 기술한다. 상기 방법은 상기 화합물의 존재 및 부재하에서, 결합체와 이들의 표적, 예컨대, 단백질의 식별에 기반하고 있다. 결합체와 이들의 표적 사이에 형성된 복합체의 검출, 및 상기 결합체의 식별은 바람직하게는 복수의 구획에서, 결합된 표적 특이적 결합체 (이는 이후 뉴클레오티드로 번역된다)의 고유의 식별 서열의 쌍을 이루는 방식의 연결(linkage)에 의해 수행된다. 복합체의 양과 결합체 및 선택적으로 포함되는 표적의 동일성의 변화는 서열분석으로 밝혀질 수 있다.

[0026] 상기 결합체가 바람직하게는 항체, 또는 암타머(aptamer)이다.

[0027] 바람직하게는, 상기 결합체가 항체 디스플레이 라이브러리 또는 항체의 라이브러리 (여기서 각 항체는 상기 고유의 뉴클레오티드 서열로 표지된다)의 멤버이다.

[0028] 바람직하게는, 상기 표적이 또한 고유의 뉴클레오티드 서열과 연관되어 있다.

[0029] 복합체 중의 결합체와 연관되어 있는 뉴클레오티드 서열을 복합체 중의 제2의 결합체와 연관되어 있는 제2의 뉴클레오티드 서열에 연결시킬 수 있다. 이 방법을 사용하여 단일 표적에 결합하는 복수의 결합체를 식별할 수 있다. 예를 들어, 표적이 단백질일 때, 상기 방법은 상기 단백질 상에 있는 상이한 에피토프에 결합하는 항체를 식별할 수 있다. 달리, 상기 표적이 단백질 복합체일 수 있으며, 상기 방법이 상기 복합체 내에 있는 상이한 단백질에 결합하는 복수의 결합체를 식별할 수 있다. 예를 들어, 상기 결합체/표적 복합체 중의 하나의 결합체와 연관되어 있는 뉴클레오티드 서열을 결합체/표적 복합체 중의 제2의 결합체와 연관되어 있는 뉴클레오티드 서열에 연결할 수 있다. 일단 결합체의 동일성이 (연결된 서열로부터) 알려지면, 표적의 구성분(component)을 식별할 수 있으며, 따라서, 예를 들면, 자연적으로 상호반응하는 표적 중의 단백질을 식별할 수 있다. 예를 들어, 결합체가 알려져 있는 결합 특성을 갖는 항체인 경우, 상기 항체에 의해 결합된 단백질이 식별될 수 있다. 따라서, 표적 내의 단백질의 동일성이 식별될 수 있다. 이는 샘플 내의 단백질-단백질 상호반응을 검출하고 확인할 수 있도록 한다. 또한, 일단 단백질-단백질 상호반응이 식별된다면, 상기 방법을 사용하여 상기 상호반응에 대한 화합물의 효과를 모니터할 수 있다.

[0030] 달리, 결합체/표적 복합체 중의 결합체와 연관되어 있는 뉴클레오티드 서열을 결합체/표적 복합체 내의 표적과 연관되어 있는 뉴클레오티드 서열에 연결할 수 있다. 이 방법을 사용하여 어느 결합체가 어느 표적과 상호반응하는지 식별할 수 있다. 예를 들어, 결합체 라이브러리의 멤버가 알려져 있는 표적과 복합체를 형성할 수 있는지 식별하는데 사용될 수 있다. 이 정보를 사용하여 결합 특성 정보를 얻기 위하여 결합체 라이브러리의 멤버를 특성화할 수 있다.

[0031] 바람직하게는, 상기 랜덤하게 쌍을 이룬, 연결된 핵산 생성물의 생산은 동일하거나 동일하지 않은 앰플리콘(amplicon)을 증폭시키기 위하여 적어도 2쌍의 PCR 프라이머를 이용하는 것을 포함하며; 여기서 5' 말단에서 상기 PCR 프라이머는 서열 태그(tag)를 가지며 이때 태깅된(tagged) 프라이머를 사용한 증폭으로 랜덤하게 쌍을 이룬, 연결된 핵산 생성물이 생성된다. 더욱 바람직하게는, 증폭이 애벌전 PCR 증폭이며 상기 앰플리콘 및 랜덤하게 쌍을 이룬, 연결된 핵산 생성물의 생산은 평행 공정(parallel process)이다.

[0032] 바람직하게는, 상기 결합된 증폭 생성물의 서열분석이 고도의 평행 서열분석법이다.

[0033] 본 발명의 방법을 사용하여 결합체와 표적 간의 상호반응 또는 2개 이상의 표적 분자 간의 상호반응에 대한 화합물의 효과를 조사할 수 있다. 결합체와 표적의 접촉 단계를 화합물의 존재 및 부재하에서 수행할 수 있으며, 결과를 비교하여 상기 화합물이 결합체와 표적 간 또는 표적 분자 간의 결합 상호반응에 영향을 주는지 결정한다. 이 방법을 이용하여 내과적 질환 및 병증을 치료하는데 사용될 수 있는 잠재적 약학적 화합물을 식별할 수 있다.

- [0034] 본 발명은 또한 본 발명의 방법을 수행하기 위한 키트를 제공하며 이는
- [0035] (i) 결합체 라이브러리 (여기서 상기 결합체 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다); 및
- [0036] (ii) 상기 결합체와 연관되어 있는 뉴클레오티드 서열을 연결하기 위한, 적어도 2쌍의 프라이머 세트; 및 선택적으로 사용설명서를 포함한다.
- [0037] 상기 키트는 상기 라이브러리의 각 멤버가 고유의 뉴클레오티드 서열과 연관되어 있는 단백질 디스플레이 라이브러리를 추가로 포함할 수 있다.

도면의 간단한 설명

- [0038] 도 1은 본 발명의 방법과 조성물을 사용하여 단백질-단백질 상호반응을 검출하기 위한 분석법의 일반적인 원리를 도시한 것이다. 예정된 결합 특성 정보를 갖는 항체 라이브러리의 파아지를 사용하여 상호반응하는 단백질의 결합 정보를 밝힌다. 결합 정보는 PCR 다이머화(dimerization)에 의해 복수의 구획에서 결정된다. 상기 파아지는 공정 중에 용해되어 이들의 고유의 DNA가 방출된다. 이들 고유의 서열은 범용 프라이머를 사용하여 증폭되고 다이머화된다. 결합 정보를 코딩하는 상기 다이머화된 생성물을 차세대 서열분석법(NGS)으로 서열화하고 결합된 단백질 동일성을 인식된 결합 표적의 동일성을 포함하여 알려져 있는 결합 특성을 갖는, 특이적인 파아지의 검출을 기본으로 하여 결정된다.
- 도 2는 본 발명의 방법 및 조성물을 사용하여 인식된 결합 표적의 동일성을 포함하여, 결합 특성 정보에 대한 결합체 핵산 표지 동일성을 연관시키기 위한 - 결합체의 특성화를 위한 검정법의 일반적인 원리를 도시한 것이다. cDNA 디스플레이 및 항체 디스플레이 파아지는 서로 결합한다. 결합 정보는 복수의 구획에서 PCR 다이머화에 의해 결정된다. 상기 파아지는 공정 중에 용해되어 이들의 고유의 DNA가 방출된다. 이들 고유의 서열은 범용 프라이머를 사용하여 증폭되고 다이머화된다. 결합 정보를 코딩하는 상기 다이머화된 생성물을 차세대 서열분석법(NGS)으로 서열화하고 cDNA의 단백질 동일성은 데이터베이스 검색에 의해 상기 서열로부터 결정되며 결합체 핵산 표지 동일성과 연관시킨다.

발명을 실시하기 위한 구체적인 내용

- [0039] 이상적인 방법은 시간 및 비용 효과적인 방식으로 상호반응체로부터 정보를 포착하여, 랜덤 샘플링 및 높은 리던던시(redundancy)의 샘플링을 가능하도록 한다. 상기 방법은 정량적인 상호반응 데이터, 심지어는 큰, 다중-유니트 단백질 복합체에 대한 상호반응 데이터의 포괄적인 커버리지를 제공한다. 상기 데이터는 원래의 세포 컨텍스트(original cellular context)에서 수득되며 따라서 천연의 단백질-단백질 상호반응(native protein-protein interaction)을 측정할 수 있으며, 이를 사용하여 동적 상호반응을 검출할 수 있다. 상기 방법은 비-특이적이고, 돌발적으로 상호반응하는 단백질을 검출하는 것과 같은, 랜덤한 변수의 효과를 억제한다. 이는 또한 원래의 단백질-단백질 상호반응을 제외한, 검출 원리에 연루되는 임의의 결합 반응 연관 효과, 예를 들면 자가-결합 또는 특이적 결합인, 변수의 효과를 없애준다.
- [0040] 본 발명의 하나의 실시양태가 도 1에 요약되어 있으며 상기 검정 시스템의 상이한 성분이 하기에서 상세하게 기술된다. 본 발명의 추가의 실시양태가 도 2에 요약되어 있으며 상기 시스템의 추가적인 성분이 하기에서 상세하게 기술된다.
- [0041] 본 발명은 결합체와 표적 간의 결합 상호반응을 측정하기 위한 방법을 제공하며, 이는
- [0042] a) 결합체 라이브러리와 표적을 접촉시켜 결합체/표적 복합체가 형성되도록하는 단계 (여기서 상기 결합체 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다);
- [0043] b) 상기 결합체/표적 복합체를 분리시키는 단계;
- [0044] c) 결합체/표적 복합체에서 결합체와 연관되어 있는 뉴클레오티드 서열을 연결하여 연결된 뉴클레오티드 서열을 형성시키는 단계;
- [0045] d) 상기 연결된 뉴클레오티드 서열로부터 복합체에 존재하는 결합체를 식별하는 단계를 포함한다.
- [0046] 상기 방법은 생체외(ex-vivo), 생체내(in-vivo) 또는 시험관내(in vitro)에서 수행될 수 있다.
- [0047] 결합체

- [0048] 바람직하게는, 상기 결합체가 항체, 암타머이거나, 조작된 단백질 스캐폴드에 기반한 것이다. 달리, 상기 결합체가 화합물일 수 있다. 결합체는 항체 디스플레이 라이브러리 또는 각 항체가 고유의 뉴클레오티드 서열로 표지되어 있는 항체의 라이브러리의 멤버일 수 있다. 상기 방법은 결합체로서 디스플레이된 항체 화합물을 사용할 수 있으며, 여기서 결합 특성, 예를 들면, 결합체가 결합하는 표적이 알려져 있고 복수의 디스플레이된 항체 화합물과 연관있는 고유의 뉴클레오티드 서열이 결정되며 상기 결합 특성과 고유의 뉴클레오티드 서열이 서로 관련되어 있다. 따라서, 본 발명은 결합 특성을 측정하고 복수의 디스플레이된 항체 화합물의 식별가능한 고유의 뉴클레오티드 서열에 이들을 관련시키는 방법을 제공한다. 이는 결합 특성 정보를 제공한다.
- [0049] 본 발명에서 사용되는 결합체가 항체일 수 있다. 본 발명에서 사용되는 바와 같은 용어 "항체(antibody)"는 천연의 것이거나 부분적 또는 전체적으로 합성에 의해 생산되는, 면역글로불린 분자 및 면역글로불린 분자의 면역학적으로 활성인 부분, 즉, 항원에 특이적으로 결합하는 항원 결합 부위를 함유하는 분자를 언급한다. 상기 용어 "항체"는 면역글로불린 결합 도메인을 포함하는 임의의 폴리펩티드를 포함한, 항체 단편, 항체의 유도체, 기능성 등가물 및 동족체, 인간화된 항체(humanised antibodies)를 포함하며, 천연의 것이거나 전체적 또는 부분적으로 합성이며 결합 도메인을 가지거나 이에 동족성(homologous)인 임의의 폴리펩티드 또는 단백질이거나, 항체 결합 도메인이다. 따라서, 다른 폴리펩티드에 융합되어 있는, 면역글로불린 결합 도메인을 포함하는 키메릭 분자 또는 등가물이 포함된다. 키메릭 항체의 클로닝(cloning) 및 발현은 EP-A-0120694 및 EP-A-0125023에 기재되어 있다. 항체의 예로 면역글로불린 이소타입(isotype) (예를 들면, IgG, IgE, IgM, IgD 및 IgA) 및 이들의 이소타입의 서브클래스; 항원 결합 도메인을 포함하는 단편, 예컨대 Fab, scFv, Fv, dAb, Fd; 및 디아바디(diabody)가 있다. 항체는 폴리클로날 또는 모노클로날일 수 있다.
- [0050] 상보성 결정 영역(CDRs)은 B-세포에 의해 발생되는, 면역글로불린(항체) 중의 가변쇄의 일부이며, 여기서 이들 분자는 이들의 특이적 항원에 결합한다. 상기 분자의 가변성이 가장 큰 부분으로서, CDRs는 면역글로불린에 의해 생성되는 항원 특이성의 다양성에 있어서 중요하다. 면역글로불린의 가변 도메인의 아미노산 서열 상에, 비-연속적으로 배열되는, 3개의 CDR들(CDR1, CDR2 및 CDR3)이 있다. 면역글로불린이 전형적으로 2개의 가변 도메인으로 (2개의 상이한 폴리펩티드 쇄, 중쇄 및 경쇄 상에) 구성되기 때문에, 항원과 접합적으로 접촉할 수 있는 각 항원 수용체에 대해 6개의 CDR이 있다.
- [0051] 항체 전체의 단편이 결합 항원의 기능을 수행할 수 있는 것으로 나타났다. 결합 단편의 예로 (i) VL, VH, CL 및 CH1 도메인으로 이루어진 Fab 단편; (ii) VH 및 CH1 도메인으로 이루어진 Fd 단편; (iii) 단일 항체의 VL 및 VH 도메인으로 이루어진 Fv 단편; (iv) VH 도메인으로 이루어진 dAb 단편 (Ward, E.S. et al., *Nature* 341:544-546 (1989)); (v) 단리된 CDR 영역; (vi) 2개의 연결된 Fab 단편을 포함하는 2가 단편인, F(ab')2 단편; (vii) VH 도메인과 VL 도메인이 연합하여 항원 결합 부위를 형성하도록 하는 펩티드 링커에 의해 2개의 도메인이 연결되어 있는, 단일쇄 Fv 분자(scFv) (Bird et al., *Science* 242:423-426 (1988); Huston et al., *PNAS USA* 85:5879-5883 (1988)); (viii) 이중특이적 단일쇄 Fv 다이머(PCT/US92/09965) 및 (ix) 유전자 융합에 의해 구축된 다가 또는 다중특이적 단편인, "디아바디(diabodies)" (WO94/13804; P. Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90: 6444-6448 (1993))가 있다.
- [0052] "항원 결합 도메인(antigen binding domain)"은 항원에 특이적으로 결합하며 항원의 일부 또는 전체에 상보성인 영역을 포함하는 항체의 일부이다. 항원이 큰 경우, 항체는 항원의 특정 부분에만 결합할 수 있으며, 상기 부분을 에피토프(epitope)라 부른다. 항원 결합 도메인이 1개 이상의 항체 가변 도메인에 의해 제공될 수 있다. 항원 결합 도메인은 항체 경쇄 가변 영역(VL) 및 항체 중쇄 가변 영역(VH)을 포함할 수 있다.
- [0053] 달리, 결합체가 조작된 단백질 스캐폴드(engineered protein scaffold))에 기반할 수 있다. 단백질 스캐폴드는 관심의 대상이 되는 표적 분자에 대해 결합 부위를 제공하도록 변형시킨, 안정하고, 가용성인 천연 단백질 구조물로부터 유도된다. 조작된 단백질 스캐폴드의 예로, α-헬릭스 중 2개에 결합 인터페이스를 제공하는 스테필로코커스(staphylococcal) 단백질 A의 Z-도메인에 기반한 아피바디(affibody) (Nygren, P. A. (2008). *FEBS J* 275(11): 2668-76); 베타-배렐 폴드의 개방 말단에 작은 리간드용 결합 부위를 포함하는, 리포칼린(lipocalin)으로부터 유도된 안티칼린(anticalin) (Skerra, A. (2008) *FEBS J* 275(11): 2677-83), 나노바디, 및 DARPin^o 있으나, 이들로 제한되는 것은 아니다.
- [0054] 조작된 단백질 스캐폴드는 전형적으로 항체와 동일한 항원성 단백질에 결합하도록 표적화된다. 짧은 펩티드를 또한 사용하여 표적 단백질에 결합할 수 있다. 파일로머(phylomer)는 세균 게놈으로부터 유도된 천연 구조 펩티드이다. 그러한 펩티드는 단백질 구조 폴드의 다양한 어레이를 나타내며 생체내에서 단백질-단백질 상호반응을 억제/방해하기 위하여 사용될 수 있다 (Watt, P. M. (2006). *Nat Biotechnol* 24(2): 177-83).

[0055]

달리, 결합체가 압타머(aptamer)일 수 있다. 압타머는 합성 올리고뉴클레오티드 (DNA 또는 RNA)로 형태 상보성과 비-공유결합성 화학 결합의 조합을 통하여 높은 친화성과 특이성을 갖고 표적 분자를 인식한다 (*Blank & Blind, Current Opin. Chem. Biol.*, 2005, 9:336-342). 이러한 인공 리간드는 시험관내에서 수득하기가 정말로 용이하며 겨우 이온 (예, Pb²⁺, *Liu & Lu*, 2003. *J Am Chem Soc.*, 125, 6642-6643)에서부터 뉴클레오티드, 작은 분자, 단백질, 바이러스, 및 세포에서 전체 유기체(*Menger et al.*, 2006. *Handbook of Experimental Pharmacology*, 359-373)에 이르기까지, 상이한 분자 부류를 매우 다양하게 인식하도록 개발될 수 있다. 결합 친화성이 높은 압타머가 테오필린 (*Jenison et al.*, 1994. *Science*, 263, 1425-1429), L-아르기닌 (*Geiger et al.*, 1996. *Nucl. Acids Res.*, 24, 1029-1036), 모에노마이신(*moenomycin*; *Schuerer et al.*, 2001. *Bioorg. Med. Chem.*, 92, 2557-2563), 17b-에스트라디올 (*Kim et al.*, 2007. *Biosens. Bioelectron.*, 22, 2525-2531)과 같은 저분자량 분자의 검출을 위하여, 잘 알려져 있는 SELEX 방법 (*Ellington & Szostak*, 1990. *Nature*, 346, 818-822)을 통하여 선택되었지만 다른 것들 중에서도 (검토를 위하여 문헌: *Tombelli et al.*, 2007. *Biomol Eng.*, 24, 191-200를 참조), 트롬빈 (트롬빈-결합 압타머: 5'-GGTTGGTGTGGTTGG-3') (*Baldrich et al.*, *Anal Chem.* 2004, 76, 23, 7053-63), 콜레라 독소 또는 HIV-1 tat 단백질과 같이 더 큰 분자용으로도 선택되었다. 상기 언급된 압타머 중 일부는 마이크로플레이트 상에서 또는 바이오센서 변환기 (QCM, SPR)의 표면 상에서의 ELISA-유사 검정법에 사용되고 있다. 압타머-변형된 AuNP 측색 시스템이 또한 샌드위치-기반 검정법 (*Huang et al.*, 2005, 77, 5735-5741)에서 단백질 PDGF의 측정용으로 개발되었다.

[0056]

결합체는 디스플레이된 결합체 라이브러리, 예를 들면, 세균성 디스플레이, mRNA 디스플레이, 박테리오파아지 디스플레이, 압타머, 리보솜 디스플레이 또는 효모 디스플레이 라이브러리와 같은 라이브러리의 일부일 수 있다. 바람직하게는, 상기 디스플레이된 결합체 라이브러리가 항체 박테리오파아지 디스플레이 라이브러리 (antibody bacteriophage display library)이다. 상기 라이브러리는 라이브러리가 표적 샘플 내에서 관심의 대상인 표적의 적어도 75%에 결합할 것으로 예측되는 복수의 결합 멤버로 이루어지기에 충분할 정도로 커야 한다. 더욱 바람직하게는, 상기 라이브러리가 샘플 내에서 표적의 적어도 80%, 85%, 90%, 95%, 97.5% 또는 99%에 결합하도록 디자인된다. 예를 들면, 결합체 라이브러리가 샘플 내에서 예측되거나 목적하는 단백질을 95% 이상의 커버리지로 단백질 또는 펩티드 서열에 결합하는 복수의 멤버를 포함한다. 그러한 라이브러리가 문헌에 공개되어 있다. 상기 라이브러리의 각 멤버는 검출가능한 핵산 동일성 표지물(nucleic acid identity label)을 가지며, 이 표지물(label)은 상기 라이브러리의 멤버 중 하나에 대해 고유한 것이 바람직하다. 바람직하게는 고유의 핵산 동일성 표지물이 연결되어 있다. "연결된(Linked)"이란 연결 과정이 적합한 검정 조건하에서 이러한 핵산 동일성 표지물의 동시-국소화를 기반으로 랜덤한 멀티미 핵산 생성물을 형성할 수 있는 잠재성을 가지고 있음을 의미한다. 바람직하게는 상기 멀티미 생성물이 다이머(dimer)이다. 상기 적합한 검정 조건에는 예를 들어, 지질에 멀전에서의 열처리에 의한, 박테리오파아지 입자의 정형(dismantling), 바람직하게는 별개의 구획에서의 정형이 포함되며, 고유 서열의 특이적 컨센서스 증폭(consensus amplification)으로 연결가능한 앰플리콘(amplicon)이 생산된다. 연결가능한 앰플리콘, 예를 들면, 결합 디스플레이 특이적 핵산 도메인을 결합시켜 동일성 표지물의 동시-국소화 정보를 코딩하는, 연결된 동일성 표지물을 형성시킨다. 바람직하게는, 고유 서열이 결합 디스플레이 특이적 핵산 도메인, 예를 들면, 1개 이상의 CDR 영역을 코딩하는 서열이다. 결합 반응(joining reaction)은 다른 기법에 기반하는 증폭일 수 있거나 다른 기법을 포함한다. 증폭 기반 결합은 2개 이상의 증폭 프라이머 쌍을 이용할 수 있으며 증폭 프라이머 쌍은 동일한 결합 능력을 가지지만, 상보적 5' 태그 또는 다이머 링커 서열을 가지고 있어 폴리머라제 확장가능한 핵산 듀플렉스(nucleic acid duplex)를 형성한다. 상기 태그 또는 다이머 링커 서열은 하나의 프라이머 쌍에 의해 증폭된 서열이 제2의 프라이머 쌍에 의해 증폭된 서열에 하이브리드화되는 것을 의미한다. 이에 의해 동일성 표지물이 연결된다.

[0057]

결합체 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되며, 고유의 뉴클레오티드 서열은 결합체를 식별하는데 사용될 수 있다. 본 발명에서 사용되는 바와 같은 "연관된(Associated)"이란 복합체 중의 결합체의 존재가 이 방법에서 생성되는 연결된 서열 내의 핵산 서열의 존재에 의해 검출될 수 있음을 의미한다. 뉴클레오티드 서열은 결합체에 표지물로서 부착될 수 있으며, 결합체 자체의 일부, 예를 들면, 압타마일수 있거나, 결합체 내에, 예를 들면 파아지 내의 핵산으로 존재할 수 있다. 예를 들어, 상기 라이브러리의 각 멤버가 고유의 뉴클레오티드 서열로 표지될 수 있다. 본 발명에서 사용되는 바와 같은 "표지된(Labelled)"이란 라이브러리의 멤버에 부착되는 뉴클레오티드 서열을 언급한다. 항체 또는 화합물과 같은 결합체에 뉴클레오티드를 부착시키는 방법은 당해 분야에 알려져 있다. 달리, 결합체 라이브러리가 상기 기재된 바와 같은 디스플레이 라이브러리일 경우, 고유의 뉴클레오티드 서열은 1개 이상의 CDR 영역을 코딩하는 서열 또는 디스플레이된 결합 도메인일 수 있다. 예를 들어, 공지의 위치에서 파아지로 디스플레이될 아미노산 서열을 코딩하는 서열을 삽입함으로써 디스

플레이 라이브러리가 생성될 수 있다. 이후, 삽입된 서열을 증폭시킬 범용 프라이머가 사용될 수 있으며 이에 따라 결합 서열을 식별할 수 있다. 달리, 결합제가 암타머일 경우, 암타머 자체가 고유의 뉴클레오티드 서열일 수 있다.

[0058] 뉴클레오티드 서열은 올리고뉴클레오티드이며 단일사슬 또는 이중사슬인, RNA 또는 DNA를 포함할 수 있다. 결합제 또는 표적을 표지시키기 위하여 사용되는 뉴클레오티드는 일반적으로 길이가 염기 5 내지 150개, 예를 들면, 길이가 염기 10 내지 40개, 또는 20 내지 30개이다. 핵산을 형성하는 뉴클레오티드를 화학적으로 변형시켜 분자의 안정성을 증가시키거나, 이의 생체이용율을 개선하거나 이에 추가의 활성을 부여할 수 있다. 예를 들어, CH₃ 또는 할로겐, 예컨대 I, Br 또는 Cl로 6번 또는 8번 위치에서 피리미딘 염기를 변형시킬 수 있으며, 5번 위치에서 퓨린 염기를 변형시킬 수 있다. 변형 또는 피리미딘 염기는 또한 2 NH₂, O⁶-CH₃, N⁶-CH₃ 및 N²-CH₃를 포함한다. 2번 위치에서의 변형은 당 변형이며 전형적으로 NH₂, F 또는 OCH₃ 기를 포함한다. 변형은 또한 캡핑(capping)과 같은 3번 및 5번 변형을 포함할 수 있다.

[0059] 달리, 변형된 뉴클레오티드, 예컨대 모르폴리노 뉴클레오티드, 잠긴 핵산(LNA) 및 웨티드 핵산(PNA)이 사용될 수 있다. 모르폴리노 올리고뉴클레오티드는 상이한 모르폴리노 서브유니트로부터 어셈블되며, 이들 각각은 6-원 모르풀린 고리에 연결된, 4개의 유전자 염기 (아데닌, 사이토신, 구아닌, 및 티민) 중 하나를 포함한다. 상기 서브유니트는 비-이온성 포스포로디아미데이트 인터서브유니트(phosphorodiamidate intersubunit) 연결에 의해 결합되어 모르풀리노 올리고뉴클레오티드를 생산한다. LNA 모노머는 푸라노오스 고리 입체구조가 2'-0 위치를 4'-C 위치에 연결하는 메틸렌 링커에 의해 제한되는 것으로 특징된다. PNA는 골격이 당 대신 슈도웨티드인 DNA의 유사체이다.

[0060] 바람직하게는, 결합제가 1개 이상의 표적을 바람직하게는, 상이하고 분명한 친화성을 가지고 검출할 수 있다. 달리, 결합제가 상이한 에피토프 또는 결합 부위를 사용하여, 바람직하게는 상이하고 분명한 친화성을 가지고 단일 표적을 검출할 수 있다.

[0061] 항체 파아지 라이브러리의 멤버의 결합 특성이 미리 결정될 수 있다. 예를 들어, 에피토프가 파아지의 CDR 코딩되고 발현된 결합제(항체)에 의해 결합되는지 결정될 수 있다. 이 정보는 CDR을 코딩하는 고유의 뉴클레오티드 서열과 연관될 수 있다. 따라서, 파아지에 의해 발현된 항체에 의해 결합된 에피토프가 고유의 뉴클레오티드 서열의 서열로부터 식별될 수 있다. 일단 결합된 표적에 존재하는 에피토프 서열이 알려지게 되면, 단백질 또는 결합된 단백질 그룹을 식별할 수 있다.

[0062] 고유의 뉴클레오티드 서열 표지된 에피토프 또는 고유의 뉴클레오티드 서열 표지된 에피토프 라이브러리를 사용하여 항체 파아지 라이브러리의 멤버의 결합 특성을 측정할 수 있다.

표적

[0064] 본 발명에서 사용되는 바와 같은 "표적(Target)"은 결합제와 복합체를 형성하는 문자 또는 문자 그룹이다. 상기 복합체는 통상적으로 관심의 대상인 유기체의 정상적인 생리학적 상태하에서 형성된다.

[0065] 바람직하게는, 표적이 단백질을 포함한다. 더욱 바람직하게는, 표적인 단백질 샘플의 일부이다. 단백질 샘플은 단백질 디스플레이 라이브러리, 바람직하게는 상기 라이브러리의 각 멤버가 고유의 뉴클레오티드 서열과 연관되어 있는 것을 포함할 수 있다. 바람직하게는, 상기 단백질 디스플레이 라이브러리가 cDNA 파아지 디스플레이 라이브러리이다. 선택적으로, 표적이 복수의 표적, 예를 들면, 단백질 샘플 내에서 다른 표적에 가교-결합될 수 있다. 예를 들어, 샘플 내의 단백질이 샘플 내에서 1개 이상의 다른 단백질에 가교-결합될 수 있다.

[0066] 표적이 공지된 표적일 수 있다. 표적과 상호반응하는 화합물을 포함하여, 표적과 복합체를 형성하는 결합제가 식별될 수 있다. 달리, 표적이 미지의 것일 수 있으며, 본 발명의 방법을 사용하여 표적, 또는 다른 표적과 상호반응하는 복수의 표적 문자를 식별할 수 있다.

[0067] 표적이 고유의 뉴클레오티드 서열과 연관될 수 있다. "연관된(Associated)"이란 결합제/표적 복합체 내의 표적의 존재가 본 방법에 의해 생성되는 연결된 서열 내의 핵산 서열의 존재에 의해 검출될 수 있음을 의미한다. 뉴클레오티드 서열은 표지물로서 표적에 부착될 수 있거나, 표적 내에, 예를 들면, 파아지 내의 핵산과 같이 존재할 수 있다. 달리, 뉴클레오티드 서열이 표적에 결합하는 것으로 알려진 암타머의 일부일 수 있다. 결합제/표적 복합체를 암타머와 접촉시켜 존재하는 표적이 암타머를 포함하는, 고유의 뉴클레오티드 서열의 연결을 통하여 식별될 수 있도록 할 수 있다.

[0068]

본 발명의 검정법은 임의의 단백질 샘플에 적용될 수 있다. 단백질은 조직, 세포학적 표본, 체액, 세포 배양 또는 임의의 다른 단백질 복합체 함유 물질을 포함하지만, 이들로 제한되지 않는, 임의의 생물학적 표본으로부터 유래할 수 있다. 체액 샘플은 혈액, 타액, 오줌, 뇌척수액, 또는 혈청을 포함한다. 달리, 샘플이 재조합 발현 방법에 의해 생성될 수 있다. 표본으로부터 단백질의 제조는 당해 분야에 알려져 있는 표준 방법을 사용하여 수행될 수 있다. 표본은 추출 전에 화학적으로, 예를 들면, 상이한 정착성 화학물질 또는 가교결합제 (예, BS3-비스(술포숙시니미딜)수버레이트(BS3-(bis(sulfosuccinimidyl)suberate)))가 사용될 수 있다)로 처리될 수 있다. 단백질 샘플은 가교결합되거나 가교결합되지 않을 수 있다. 달리, 예를 들면, 시험관내 전사-번역 시스템에 의해, 또는 재조합 발현 시스템에 의해 단백질이 생산될 수 있다. 실험의 목적 및 조사중인 단백질-단백질 상호반응의 타입에 따라서, 단백질을 이들의 변성된 형태 또는 변성되지 않은 형태 및/또는 가교결합되거나 가교결합되지 않은 형태로 분석할 수 있다. 단백질 샘플을 복수의 조건으로 분석하여 복수의 단백질-단백질 상호반응의 정량적 결합 특성에 대한 정보를 수집할 수 있다. 예를 들어, 결합제의 농도 또는 양을 변화시켜 해리 상수와 기타 운동역학적 파라미터를 측정할 수 있다.

[0069]

단백질 혼합물을 미리 선택할 수 있다. 예를 들어, 단백질 혼합물이 특이적 단백질, 예를 들면, 특이적 세포 위치로부터, 특이적 세포 타입으로부터, 유사한 크기 또는 정전하를 갖는 단백질, 유사한 결합 특성, 유사한 서열 특성, 또는 유사한 기능을 갖는 단백질, 예를 들면 효소가 강화된 것일 수 있다 (Current Protocols in Molecular Biology (2006) 20.0.1-20.0.6 CHAPTER 20 Analysis of Protein Interactions.). 바람직하게는, 특이적 단백질이 인단백질, 막단백질 또는 천연적으로, 번역후, 인공적으로 변형된 단백질을 포함한다. 상기 방법의 단백질 혼합물 중 단백질은 변성되거나 변성되지 않고/않거나 가교결합되거나 가교결합되지 않을 것일 수 있다.

[0070]

상기 단백질이 단백질 디스플레이 라이브러리의 형태로 존재할 수 있다. 예를 들면, 세균 디스플레이, mRNA 디스플레이, 박테리오파아지 디스플레이, 및 리보솜 디스플레이와 효모 디스플레이 라이브러리를 포함한다. 바람직하게는, 단백질 디스플레이 라이브러리가 단백질 박테리오파아지 디스플레이 라이브러리, 더욱 바람직하게는 cDNA 파아지 디스플레이 라이브러리이다. 상기 라이브러리는 샘플에서 본 방법에 의해 검출될 것으로 기대되는 단백질의 적어도 70%를 커버하는 복수의 펩티드 또는 단백질 멤버로 이루어지기에 충분할 정도로 커야 한다. 더욱 바람직하게는, 상기 라이브러리가 샘플 중의 단백질 또는 펩티드의 75%, 80%, 85%, 90%, 95%, 97.5%, 99% 이상의 커버리지를 제공하기에 충분할 정도로 크다. 바람직하게는, 상기 디스플레이 라이브러리가 임의의 적합한 생물학적 개체(entity), 예를 들면 조직 샘플 또는 전체 유기체의 커버리지, 예를 들어, 임의의 적합한 생물학적 개체의 95% 이상의 단백질 커버리지를 제공한다. 그러한 라이브러리가 문헌에 공개되어 있다 (Danner S, Belasco JG. T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. Proc Natl Acad Sci U S A. 2001 Nov 6;98(23):12954-9. Epub 2001 Oct 23. PubMed PMID: 11606722; PubMed Central PMCID: PMC60806.). 상기 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되는데, 즉, 각 멤버가 고유의 검출가능한, 핵산 동일성 표지물을 가지고 있다. 바람직하게는, 상기 고유의 핵산 동일성 표지물이 연결되어 있다. "연결된(Linked)"이란 적합한 검정 조건하에서 이들 핵산 동일성 표지물의 동시-국소화를 기반으로 웬덤한 멀티머 핵산 생성물을 형성할 수 있는 잠재성을 가지고 있는 연결 과정을 의미한다. 바람직하게는, 상기 멀티머 생성물이 다이머이다. 상기 적합한 검정 조건에는 예를 들어, 지질 에멀젼에서의 열처리에 의한, 박테리오파아지 입자의 정형(dismantling), 바람직하게는 별개의 구획에서의 정형이 포함되며, 고유 서열의 특이적 컨센서스 증폭으로 연결가능한 앰플리콘이 생산된다. 연결가능한 앰플리콘, 예를 들면, 결합 디스플레이 특이적 핵산 도메인을 결합시켜 동일성 표지물의 동시-국소화 정보를 코딩하는, 연결된 동일성 표지물이 형성된다. 결합 반응(joining reaction)은 다른 기법에 기반한 증폭일 수 있으며 다른 기법을 포함할 수 있다. 증폭 기반 결합은 2개 이상의 증폭 프라이머 쌍을 이용할 수 있으며 증폭 프라이머 쌍은 동일한 결합 능력을 가지지만, 상보적 5' 태그 또는 다이머 링커 서열을 가지고 있어 폴리미라제 확장가능한 핵산 듀플렉스를 형성한다. 상기 태그 또는 다이머 링커 서열은 하나의 프라이머 쌍에 의해 증폭된 서열이 제2의 프라이머 쌍에 의해 증폭된 서열에 하이브리드화되는 것을 의미한다. 이에 의해 동일성 표지물이 연결된다.

[0071]

동일성 표지물, 즉, 결합제 라이브러리 및 표적 라이브러리에 사용되는 연관된 고유의 뉴클레오티드 서열, 예컨대 단백질 디스플레이 라이브러리 및 항체 라이브러리는 이들의 생물학적 배경에서 상이할 수 있으며, 따라서 증폭 및 결합 과정이 2개의 상이한 프라이머 쌍을 기반으로 하는데, 예를 들면, 하나의 프라이머 쌍은 cDNA 기반 동일성 표지물과 같은 표적 서열을 증폭시키고 제2의 프라이머 쌍은 동일성 표지물로서 사용되는 결합제 특이적 뉴클레오티드 서열을 증폭시킨다. 상이한 표지물을 결합시킴으로써 결합제 특이적 정보를 표적 정보, 예를 들면 디스플레이된 cDNA에 의해 코딩되는 단백질에 연결시킬 수 있도록 한다. 이런 과정의 예가 도 2에 나타나

있다.

[0072] 하나의 결합제, 바람직하게는, 디스플레이된 항체 파아지가 특이적 표적, 예를 들면, 디스플레이 단백질 표적 및 상응하는 단백질을 인식할 수 있다. 이를 특이성(specificity)이라 부른다. 달리, 복수의 결합제가 하나의 표적, 예컨대 특이적 표적, 예를 들면, 디스플레이 단백질 및 상응하는 단백질을 인식할 수 있다. 이를 리던던시(redundancy)라 부른다. 유사하게도, 하나의 결합제가 예를 들어, 단백질 입체구조 또는 단백질 서열로 인한 표적 입체구조의 유사성에 기반하여, 1개 이상의 표적, 예컨대, 단백질 종을 인식할 수 있다. 이런 현성을 교차-반응성(cross-reactivity)이라 부른다. 또한, 표적 단백질의 결합제 인식이 단백질의 입체구조 또는 이의 단백질 서열에 기반한다. 이는 이들의 반응성으로 알려져 있다. 디스플레이된 결합제와 같은 결합제의 단백질 결합 친화성은 서열분석 데이터세트의 정량적 정보로부터 계산될 수 있다. 디스플레이된 결합제의 멤버의 이미 측정된 결합 특성은 반응성 및 교차-반응성을 포함할 수 있으며 계산된 친화성과 함께 특이성과 리던던시도 포함될 수 있다.

[0073] 이러한 척도를 사용하여 검출된 표적-표적, 예를 들면, 단백질-단백질 상호반응을 계산할 수 있다. 연결된 동일성 표지물의 반응성, 특이성 및 리던던시와 동일성 및 존재비(abundance)의 투입량을 기반으로 계산함으로써, 특이적 표적-표적, 예를 들면, 단백질-단백질 상호반응을 통찰할 수 있다. 유사하게, 이러한 척도를 사용하여 교차-반응성에 의해 유발되는 계산의 불확실성을 줄일 수 있는데, 이때 디스플레이된 항체 화합물의 리던던시와 친화성이 고려된다. 변화된 농도의 결합제 및/또는 표적을 사용하여 복수의 상호반응, 예컨대 단백질-단백질 상호반응에 대한 정량적 파라미터를 계산할 수 있다. 따라서, 바람직하게는, 본 방법이 상이한 농도의 결합제 및/또는 표적을 사용하여 수행된다.

[0074] 정량적 특성의 검출로 동일성 표지물의 비-특이적 동시-국소화(non-specific co-localisation) 또는 동일한 동일성 표지물의 자가-연결(self-linking)에 의해 생산되는, 배경, 비-정보적 서열분석 판독치를 측정 및 계산할 수 있다. 하지만, 자가-연결 표지물의 검출은 데이터세트의 품질에 관한 정보를 포함한다.

[0075] 충분한 커버리지를 성취하기 위하여, 강화된 라이브러리(enriched libraries)가 사용될 수 있다. 결합제 라이브러리 및/또는 표적 라이브러리 둘 다 강화될 수 있다. 디스플레이된 표적은, 예컨대 실험 컨텍스트에서 모든 잠재적 결합 파트너를 커버할 수 있는 풍부한 단백질(enriched protein)이 될 수 있다. 유사하게, 상기 디스플레이된 결합제를 실험 컨텍스트에서 검출가능한 표적에 대해 강화된 결합특이성을 갖도록 선택할 수 있다. 예를 들어, 상기 라이브러리가 관심의 대상이 되는 특이적 표적으로 제한될 수 있다. 이는 패닝 기반 선택(panning based selection)에 의해 수행될 수 있으며, 예를 들면, 복잡성(complexity)이 조절된 표적 샘플을 고체 표면에 고정화시키고(immobilizing) 결합제 라이브러리와 접촉시켜 결합된 결합제의 세척 및 후속되는 용출(elution)에 의해 결합제 라이브러리의 결합된 결합제를 선택한다. 유사하게, 복잡성이 조절된 결합제 샘플을 고체 표면에 고정화시키고 표적 라이브러리와 접촉시켜 결합된 표적을 세척하고 후속되는 용출에 의해 표적 라이브러리의 결합된 표적을 선택한다.

[0076] 바람직하게는, 강화란 자가-결합하는 디스플레이된 결합제의 제거, 즉, 검정 조건하에서 다른 결합제에 결합하는 결합제를 제거하는 것을 의미한다. 더욱 바람직하게는, 강화가 디스플레이된 결합제 라이브러리의 복잡성을 감소시켰지만, 실험 컨텍스트에서 표적의 검출에 대해 높은 커버리지를 여전히 가지고 있음을 의미한다. 복잡성을 감소시키는 것은 라이브러리 내 멤버의 총수의 감소, 비-표적 단백질에 결합하는 멤버의 제거, 또는 관심의 대상인 표적에 결합하는 멤버들만 선택하는 것을 포함할 수 있다.

[0077] 결합제 라이브러리는 관심의 대상인 화합물과의 단백질 혼합물을 수득함으로써, 예를 들면, 공지의 단백질-단백질 상호반응을 연구하고 동적으로 이들의 상호반응을 검증하거나 작용성 또는 길항성 화합물의 효과를 테스트함으로써 강화될 수 있다. 예를 들어, 관심의 대상인 표적을 고정화시키는 것, 예를 들면, 고체 표면상의 단백질 혼합물은 결합제 라이브러리가 목적하는 표적, 예를 들면, 단백질에 결합하도록 한다. 관심의 대상인 표적에 결합된 결합제는 회석 또는 다른 수단에 의해 결합되지 않은 결합제로부터 분리될 수 있다. 결합제 라이브러리의 결합된 멤버는 용출되어 본 발명의 방법에서, 상기 기재된 바와 같이 복잡성이 감소된 라이브러리로서 사용될 수 있다.

[0078] 일반적으로 분리 방법 (예를 들면, 마이크로아레이 또는 애밀젼 기반 분리법)이 하나의 구획 내에서 각각의 복합체를 생산하는 능력이 제한되어, 연결된 식별 가능한 서열을 수득할 수 있기 때문에 강화된 결합제 라이브러리를 사용하는 것이 유리하다. 복잡성이 감소된 강화된 디스플레이 결합제 라이브러리는 분리 기반, 랜덤하게 쌍을 이룬, 결합된 종족 생성물을 수득하기 위하여 수행할 필요가 있는 수많은 분리 단계로 직접 변환(directly translatable)될 수 있다.

[0079] 결합제의 결합(binding)

본 발명의 방법은 생리학적 조건하에서 바람직하게 수행되어 상호반응을 이들의 원래의 컨텍스트로, 즉, 세포에 존재하는 것과 동일한 조건으로 검출할 수 있다. 이는 자연에서 일어나는 결합 상호반응에 대한 정보를 제공한다.

[0081] 결합제 라이브러리와 표적의 접촉 단계는 통상적으로 공지의 완충 시스템, 예를 들면 단백질-단백질 상호반응의 연구에 이미 사용되고 있는 완충 시스템 (예, TBST-완충액)에서 수행된다. 친화성에 따라서, 상기 반응을 실온 또는 4°C에서 수행할 수 있다. 재현가능한 시그널을 얻기 위하여 결합, 세척 및 검출 단계를 포함하여, 최적 시간, 최적 온도 및 기타 조건을 측정한다. 최적 조건은 당해 분야의 숙련가에 의해 결정될 수 있다.

[0082] 결합되어 디스플레이된 항체 결합제와의 단백질 복합체의 분리

[0083] 결합제/표적 복합체는 분리, 즉, 표지 서열, 즉, 연관된 고유의 뉴클레오티드 서열에 연결하기 전에 다른 복합체로부터 단리시킬 필요가 있다. 상기 분리는 당해 분야에 알려져 있는 방법에 의해 수행된다. 예를 들어, 분리가 희석, 특이적 결합, 또는 물리적 및/또는 화학적 특성에 의한 분리에 의해 수행될 수 있다. 바람직하게는, 복합체가 구획으로, 예컨대 에멀젼 액滴(emulsion droplets), 미세-공동 등(micro-cavities), 바람직하게는 확산 제한되거나 분리된 구획으로 분리된다.

[0084] 바람직하게는, 상기 분리가 그중에서도 고체 표면 결합, 희석 또는 상분리법 중 하나 이상을 포함하거나, 확산 제한되거나 분리된 구획을 제공한다. 상기 분리는 구획(compartment) 내에 결합되지 않은 결합제의 수를, 바람직하게는 평균 1개로 제한한다. 예를 들어, 구획 내에 결합되지 않은 결합제의 평균수는 1이다. 구획은 에멀젼, 또는 미세공동과 같은 개별적인 물리적 챔버 내의 각각의 액滴일 수 있다. 복합체가 물리적 또는 화학적 특성에 따라서 분리될 수 있다. 바람직하게는, 상기 희석법이 제한된 희석법(limited dilution)이다.

[0085] 구획화(compartmentalization) (예를 들면, 많은 수의 반응물의 효과적인 분리(separation) 또는 단리(isolation))는 단일 비결합 파아지의 포아송 분포(Poisson distribution) 기반 분리법에 기반하며; 예를 들면, 에멀젼 및 마이크로어레이가 당해 기술 상태에서 알려져 있는 최선의 방법이다.

[0086] 추가적인 분석 전에 결합제/표적 복합체를 충분하게 분리시킴으로써 결합제의 동시-국소화에 기반하는, 핵산 표지물의 쌍이 생성되는 환경이 제공된다. 복합체가 충분하게 분리되지 않을 경우 상이한 복합체의 멤버로부터의 핵산 표지물이 연결되게 되며 이에 따라 잘못된 정보가 제공된다. 분리는 핵산 동일성 표지물의 비-특이적 동시-국소화의 양을 감소시키며 특히 복합체 단백질 혼합물을 조사할 경우 특이적 결합 파트너의 식별을 가능하게 한다. 예를 들면, 분리는 연결이 자가-연결된 뉴클레오티드 서열만 제공하는 경우, 평균적으로, 구획당 단일 비결합된 결합제로 되며, 결과적으로 결합제 라이브러리 멤버 간의 웬덤한 연결 가능성을 감소시키게 된다. 임의의 결합제의 분포가 포아송 분포에 기반하기 때문에 결합제의 적합한 분리를 수행하는데 필수적인 척도가 계산될 수 있다. 이는 바람직하게는 하나의 구획 내의 단일 복합체로 된다. 바람직하게는, 에멀젼이 사용되며, 이는 연결된 동일성 표지물을 형성시키기 위한 증폭 및 특이적 핵산 도메인의 연결에 이용될 수 있다. 에멀젼 증폭 방법은 당해 분야의 숙련가에, 예를 들면, 문헌: Schutze et al., Anal. Biochem. 2011 March 1; 410 (1):155-7에 잘 알려져 있다.

[0087] 고정화를 위한 고체 표면

[0088] 선택적으로, 결합제/표적 복합체를 고정화시킨 다음 분리시킨다. 예를 들면, 결합제/표적 복합체를 예를 들어, 어레이의 일부로서, 표면 위에 포착시킬 수 있다. 이는 복합체 간 분리를 유지하는데 도움이 될 수 있다. 바람직하게는, 복합체의 분리를 예를 들면, 연결 단계(linking step) 중에 유지한다.

[0089] 결합제/표적 복합체를 선택적으로, 멤브레인, 예를 들면, 폴리비닐리덴 플루오라이드(PVDF) 또는 니트로셀룰로오스, 플라스틱 표면 (예, 폴리스티렌)이 포함되지만 이들로 제한되지 않는, 고체 지지체 표면에 고정화시키거나 적합한 비드 (예, 에폭시-활성화된 비드)에 공유결합적으로 커플링시킬 수 있다. 고체 지지체에 결합 또는 커플링시키는 것은 단백질에 대한 표준 방법 ("Antibodies, a Laboratory Manual." Harlow, E., and Lane, D., eds. Cold Spring Harbor Press, Cold Spring Harbor N.Y., 1988)에 의해 또는 항체를 포함하는 특이적 결합 또는 기타 특이적 결합 상호반응, 예를 들면, 비오틴-아비딘(biotin-avidin)에 의해 수행된다.

[0090] 고정화는 (a) 범용 인식된 단백질 복합체 결합 능력이 있는 고체 지지체를 수득하는 단계 (즉, 고체 지지체가 결합제 라이브러리의 모든 멤버 또는 표적 분자에 결합할 수 있다)를 포함할 수 있다. 인식된 단백질 복합체의 양과 고체 지지체상의 이용가능한 결합 부위의 수는 결합되어 인식된 단백질 복합체 간에 충분한 분리가 성취되

도록 균형잡혀야 한다. 바람직하게는, 상기 고체 지지체가 멤브레인, 플라스틱 표면, 또는 비드(bead)를 포함한다. 더욱 바람직하게는, 고체 지지체가 비드이고 분리는 평균적으로 1개의 결합체/표적 복합체가 1개의 비드에 결합되도록 수행된다.

[0091] 더욱 바람직하게는, 범용 인식된 단백질 복합체 결합 능력이 항-박테리오파아지 항체에 의해 제공된다. 바람직하게는, 분리가 반응 챔버에서의 물리적 분리 또는 에멀젼 내 액적에서의 물리-화학적 분리를 포함한다.

[0092] 바람직하게는, 상기 랜덤한, 쌍을 이룬, 연결된 핵산 생성물의 생산이 적어도 2쌍의 PCR 프라이머를 사용하여 동일하거나 동일하지 않은 앰플리콘을 증폭시키는 것을 포함하며; 여기서 5' 말단에서의 PCR 프라이머는 서열 태그를 가지는데, 이때, 태깅된 프라이머를 사용한 증폭으로 랜덤한, 쌍을 이룬, 연결된 핵산 생성물이 생성된다. 더욱 바람직하게는, 증폭이 에멀젼 PCR 증폭이며 상기 앰플리콘과 랜덤한, 쌍을 이룬, 연결된 핵산 생성물의 생산은 평행 공정이다. 바람직하게는, 상기 결합된 증폭 생성물의 상기 서열분석은 고도로 평행인 서열분석법(hightly parallel sequencing method)이다. 분리시킴으로써 연결된 동일성 표지물이 특이적으로 형성되도록 하며, 이때 분리되거나 확산 제한된 구획 (예를 들면, 고체 표면, 에멀젼 중)에 동시-국소화된 분자는 오로지 결합된 결합체의 동일성 표지물 서열을 특이적으로 연결하는 경향을 갖는다. 그러나, 결합되어 있지 않지만, 돌발적으로 동시-국소화된 결합체의 동일성 표지물이 또한 연결될 수 있다. 유사하게, 복합체 내에 결합되어 있지 않은 결합체로부터 동일성 표지물이 연결될 수 있다. 연결 과정은 예를 들어, 결합체가 구획에 하나만 존재하는 경우, 동일한 동일성 표지물을 2개 갖는 연결된 동일성 표지물을 제공할 수 있다. 유사하게, 표적을 인식하는 결합체가 1개 이상인 경우, 상이한 연결된 동일성 표지물이 동일한 결합 특이성을 갖고 생산될 수 있다. 연결 단계는 존재하는 결합체 라이브러리의 멤버 1개 이상의 동일성 표지물을 결합시킨다. 또한, 연결 단계는 표적과 연관되어 있는 핵산 서열에 결합체 라이브러리의 멤버의 동일성 표지물을 결합시킬 수 있다. 연결 과정은 예를 들면, 서로 하이브리드화시킴으로써 이들의 밀접한 근접성(proximity)으로 인하여 서로 상호반응하는 핵산에 의존하지 않는다. 연결 과정을 사용함으로써, 예를 들면, 본 명세서에 기재된 바와 같은 증폭 방법에 의해 서열이 함께 결합된다. US7306904에 기재된 것들과 같은, 밀접한 물리적 근접성 라이게이션 어세이(ligation assay)에 의존하는 대신, 이러한 방법들을 사용함으로써 다중 상호반응이 평행하게 검출되도록 한다. 상기 표지물은 연결하여 검출되도록 밀접하게 근접하여 존재할 필요가 없다. 이들은 단순히 동일한 구획 내에 있을 필요가 있다.

뉴클레오티드 서열의 연결(linkage)

*복합체의 멤버와 연관되어 있는 고유의 뉴클레오티드 서열을 다음과 같은 방법으로 연결할 수 있으며, 이 방법은

[0095] (i) 결합체와 연관되어 있는 뉴클레오티드 서열을 증폭시키고, 선택적으로, 표적과 연관되어 있는 서열이 존재할 경우, 적어도 2쌍의 PCR 프라이머를 사용하여 적어도 2세트의 앰플리콘 (이때, 프라이머는 제1 세트의 앰플리콘이 제2 세트의 앰플리콘 내의 서열에 상보적인 서열을 포함하도록 디자인된다)을 생산하는 단계;

[0096] (ii) 적어도 2세트의 앰플리콘을 어닐링(annealing)시키는 단계; 및

[0097] (iii) 증폭 반응을 수행하여 연결된 뉴클레오티드 서열을 생산하는 단계를 포함한다.

[0098] 단계 (i)-(iii)은 순서대로 또는 동시에 수행될 수 있다.

[0099] 각 쌍의 프라이머는 정방향 프라이머와 역방향 프라이머를 포함한다. 이들 프라이머의 서열은 이들이 식별가능한 서열의 증폭을 허용하도록 디자인된다. 바람직하게는, 이들이 범용 프라이머, 즉, 이들이 라이브러리 내의 모든 식별가능한 서열, 예를 들면, 결합체 라이브러리, 또는 표적 라이브러리의 모든 멤버에 결합한다. 증폭되는 프라이머 사이의 서열은 라이브러리의 멤버 하나에 대해 고유하여, 식별을 가능케 한다. 바람직하게는, PCR 프라이머 쌍들이 다이머화된 연결된 핵산 서열을 생산하도록 디자인되거나, 달리 멀티머(multimer) 연결된 핵산 서열이 생산된다. 이는 다이머 또는 멀티머의 멤버 앰플리콘을 증폭시키는 프라이머 적어도 2개의 5' 말단에 다이머 링커 서열을 가짐으로써 성취된다. 따라서, 이들 앰플리콘은 이들의 3' 말단에서 부분적으로 중첩되며 폴리머라제로 확장 가능한 하이브리드화 생성물을 형성할 수 있다 (예를 들면, 도 2에 나타낸 바와 같음).

[0100] 바람직하게는, 증폭이 결합체와 연관되어 있는 뉴클레오티드 서열, 또는 결합체와 표적의 적합한 분리 단계가 있는 에멀젼 PCR 증폭이다.

[0101] 선택적으로, 복합체 1개 이상으로부터의, 연결된 뉴클레오티드 서열은 연결된 결합체 및/또는 존재하는 표적을 식별하기 전에 합한다. 결합체(들) 및/또는 표적(들)의 동일성을 예를 들면, 연결된 뉴클레오티드 서열을 서열 분석함으로써 측정될 수 있다. 이는 고도로 평행인 시스템을 사용하여 수행될 수 있다. 연결된 서열을 합함으로

써 연결된 서열을 모두 식별하기 위하여 단일 반응을 수행할 수 있다. 예를 들면, 연결된 뉴클레오티드는 모두 단일 반응으로 서열분석될 수 있다. 연결된 서열의 상대적 존재비를 측정하기 위하여, 상기 연결된 서열을 정량적으로 측정할 수 있다.

[0102] 상기 연결된 핵산 생성물의 정보의 측정

결합체의 예정된 결합 특성을 기본으로, 동시-국소화 정보가 상기 연결된 동일성 표지물로부터 추론될 수 있다. 복수의 연결된 동일성 표지물, 즉, 결합체 및 선택적으로 표적과 연관되어 있는 고유의 뉴클레오티드 서열은 서열분석 판독치의 카운트 형태로, 결합체/표적 복합체의 멤버의 동일성 정보와 이들의 상대적인 존재비를 제공한다. 다른 정보도 추론할 수 있는데, 이때, 모든 연결된 동일성 표지물의 정도가 고려된다. 예를 들면, 결합체의 예정된 친화성과 상대적 친화성의 비교, 상이한 결합체에 대해 계산된 상대적 존재비의 비교, 결합 및 비결합 표적 또는 단백질 비율의 측정이 포함되지만 이들로 제한되는 것은 아니다.

항체 결합체의 CDR 영역 (CDR1, CDR2 및 CDR3)의 멀티머 PCR 연결을 기본으로, 모든 CDR 영역의 서열을 NGS 서열분석법을 사용하여 결정할 수 있다. 하나의 바람직한 실시양태로, 결합체의 예정된 결합 특성이 바람직하게는, 단일 CDR 영역 서열 동일성에 기반으로 하기 때문에, 항체 결합체의 전체 서열 정보가 이의 예정된 결합 특성과 연관될 수 있다.

[0105] 복수의 단백질-단백질 상호반응에 대한 화합물의 효과의 측정

본 발명의 방법은 결합체 라이브리리를 화합물의 존재 또는 부재하에서 표적과 접촉시켜; 상기 화합물이 결합 상호반응에 영향을 주는지 측정하는 단계를 포함할 수 있다. 이 방법을 사용하여 복수의 결합 상호반응, 예를 들어, 단백질-단백질 상호반응에 대한 화합물 또는 기타 화학 잔기(moiety)의 효과를 측정할 수 있다. 바람직하게는, 이 방법을 사용하여 특정 단백질-단백질 또는 기타 검출가능한 상호반응을 촉진하거나 방해할 수 있는 화합물 또는 화학 잔기를 검출할 수 있으며, 여기서 상기 화합물 또는 화학 잔기는 약물로서 작용하거나 그러한 상황의 병리학적 결과를 제거하거나 억제한다. 바람직하게는, 그러한 약물이 암, 감염성 질환, 자가면역질환 및 기타병증(을 포함하나 이들로 제한되지 않음)을 포함한 상이한 질환을 치료하는데 사용될 수 있다.

본 발명에서 사용되는 바와 같은 화합물은 공유결합에 의해 연결된 원자 2개 이상을 언급한다. 화학 잔기는 화합물의 일부로, 기능성 그룹을 형성한다. 상기 화합물은 공지되어 있는 약제일 수 있다.

본 발명의 방법은 결합체 라이브리리를 화합물의 존재 또는 부재하에서 표적과 접촉시켜; 상기 화합물이 결합 상호반응에 영향을 주는지 측정하는 단계를 포함할 수 있다.

추가의 실시양태로, 본 발명은 상호반응체 데이터를 측정하는 방법을 제공한다. 복수의 구획에서, 결합체, 예컨대 디스플레이된 항체 결합체를 이들의 결합 특성에 따라서 동시-국소화시키는데, 즉, 2개 이상의 결합체가 동일한 표적에 결합하거나 자체적으로 상호반응하거나 서로 결합되어 있는 표적에 결합하기 때문에 이들은 하나의 구획 내에 존재한다. 이들의 식별가능한 서열, 즉, 고유의 뉴클레오티드 서열이 연결되어 연결된 동일성의 형태로 동시-국소화 정보를 가지게 된다. 이런 결합 특성 정보와 동시-국소화 정보를 기본으로 하여, 단백질-단백질 상호반응 및 단백질의 동일성 둘 다 측정될 수 있다.

본 발명은 또한 수득한 표본 중의 단백질-단백질 상호반응을 측정하는 방법을 기술하고 있으며, 이 방법은:

[0111] (a) 단백질 혼합물을 수득하는 단계;

[0112] (b) 결합체 라이브리리와 상기 단백질 혼합물을 접촉시켜 결합체-단백질 복합체를 형성시키는 단계 (여기서, 상기 결합체의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다);

[0113] (c) 상기 결합체-단백질 복합체를 분리시키는 단계;

[0114] (d) 선택적으로, 상기 결합체 단백질 복합체를 고체 표면에 고정화시키는 단계;

[0115] (e) 상기 결합체-단백질 복합체 내에서 상기 결합체의 고유의 뉴클레오티드 서열을 검출하고 상기 결합체의 고유의 뉴클레오티드 서열을 연결하여 상기 분리에 의해 제공되는 동시-국소화를 기본으로 연결된 핵산 생성물을 제공하는 단계;

[0116] (f) 선택적으로, 상기 연결된 핵산 생성물을 합하는 단계; 및

[0117] (g) 상기 디스플레이된 항체 결합체의 단백질 결합 특성에 상응하는 상기 연결된 핵산 생성물을 서열분석하는 단계를 포함한다. 상기 연결된 핵산 생성물의 서열을 사용하여 상기 표본 내 상기 단백질 혼합물에서 단백질-단

백질 상호반응의 존재를 추론한다. 고유의 뉴클레오티드 서열은 존재하는 항체, 및 상기 디스플레이된 항체 제제의 상응하는 단백질 결합 특성이 식별되도록 한다.

[0118] 단백질-단백질 상호반응 데이터는 배경 측정 및 단백질-단백질 상호반응의 수준의 빼기(subtraction), 분명한 상대적 친화성 및 단백질과 단백질-단백질 상호반응의 상대적 존재비의 측정을 포함한, 통계적 수단에 의해 검증될 수 있다.

[0119] 본 발명은 또한 표본에 존재하는 단백질-단백질 상호반응에 대한 화합물의 효과를 측정하는 방법에 관한 것으로, 이 방법은:

[0120] (a) 화합물의 존재 및 부재하에서 단백질 혼합물을 수득하는 단계;

[0121] (b) 결합체 라이브러리와 상기 단백질 혼합물을 접촉시켜 결합체-단백질 복합체를 형성시키는 단계 (여기서 결합체 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다);

[0122] (c) 상기 결합체-단백질 복합체를 분리시키는 단계;

[0123] (d) 선택적으로, 상기 결합체 단백질 복합체를 고체 표면에 고정화시키는 단계;

[0124] (e) 상기 결합체-단백질 복합체 내에서 상기 결합체와 연관되어 있는 상기 고유의 뉴클레오티드 서열을 검출하고 상기 결합체와 연관되어 있는 상기 고유의 뉴클레오티드 서열을 연결하여 상기 분리에 의해 제공되는 동시-국소화를 기본으로 하는, 랜덤하게 쌍을 이룬, 연결된 핵산 생성물을 제공하는 단계;

[0125] (f) 선택적으로, 상기 연결된 핵산 생성물을 합하는 단계; 및

[0126] (g) 상기 연결된 핵산 생성물을 서열분석하는 단계;

[0127] (h) 화합물의 존재 및 부재하에서 수득한 정보를 비교하는 단계를 포함한다.

[0128] 상기 결합체의 고유의 뉴클레오티드 서열은 존재하는 결합체를 식별할 수 있도록 하여, 상기 결합체의 예정된 단백질 결합 특성을 사용하여 상기 표본 중 상기 단백질 혼합물에서 단백질-단백질 상호반응의 존재를 추론한다.

[0129] 바람직하게는, 상기 화합물의 효과의 측정은 고효율 실험 셋업을 사용하여 수행한다.

[0130] 본 발명의 추가의 실시양태로, 결합체 라이브러리의 멤버의 단백질 결합 특성을 측정하는 방법이 제공되며, 이 방법은:

[0131] (a) 디스플레이된 단백질 라이브러리를 수득하는 단계 (여기서 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다);

[0132] (b) 결합체 라이브러리와 상기 디스플레이된 단백질 라이브러리를 접촉시켜 결합체/단백질 복합체가 형성되도록 하는 단계 (여기서 상기 결합체 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다);

[0133] (c) 상기 결합체/단백질 복합체를 분리시키는 단계;

[0134] (d) 선택적으로, 상기 분리된 결합체/단백질 복합체를 고체 표면에 고정화시키는 단계;

[0135] (e) 상기 결합체 및 단백질과 연관되어 있는 고유의 뉴클레오티드 서열을 연결하여 연결된 핵산 생성물을 생산하는 단계; (선택적으로 상기 복합체의 분리를 유지하면서);

[0136] (f) 선택적으로 상기 연결된 핵산 생성물을 합하는 단계;

[0137] (g) 상기 연결된 핵산 생성물의 존재를 측정하는 단계를 포함한다.

[0138] 상기 결합체 라이브러리 멤버의 단백질 결합 특성은 서열 연결된 핵산 생성물 내에서 상기 정보로부터 측정될 수 있다. 디스플레이된 결합체 라이브러리의 멤버 및 단백질과 연관되어 있는 연결된 고유의 뉴클레오티드 서열이 검출된다는 것은 디스플레이된 단백질 라이브러리의 특정 멤버를 인식하고 결합하는 것을 디스플레이한다. 상기 서열은 라이브러리의 어느 멤버가 어느 단백질에 결합하는지에 대한 정보를 제공할 수 있다.

[0139] 복수의 상기 결합체의 단백질 결합 특성을 계산할 수 있다. 상기 결합체 라이브러리의 모든 멤버에 대한 모든 결합 특성 정보를 결합 정보로서 합할 수 있다.

[0140] 상기 결합체 및 상기 단백질과 연관되어 있는 상기 고유의 뉴클레오티드 서열을 연결함으로써 상기 분리에 의해

제공되는 동시-국소화를 기본으로 랜덤하게 쌍을 이룬, 연결된 핵산 생성물이 제공된다.

[0141] 바람직하게는, 상기 결합체 라이브러리 및/또는 표적 (예, 단백질 샘플)을 상기 복수의 단백질-단백질 상호반응의 해리 상수를 포함한 정량적 결합 정보의 계산을 허용하는 복수의 측정에 변화된 농도로 사용한다. 이러한 측정치를 계산하는 방법은 당해 분야의 숙련가에게 알려져 있다.

[0142] 본 발명은 또한 본 발명의 방법을 수행하는 위한 키트에 관한 것이다. 상기 키트는

[0143] a) 결합체 라이브러리 (여기서 상기 결합체 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다); 및

[0144] b) 상기 결합체와 연관되어 있는 뉴클레오티드 서열을 연결하기 위한 적어도 2쌍의 프라이머 세트; 및 선택적으로 사용 설명서를 포함한다.

[0145] 상기 키트는 또한 단백질-단백질 상호반응을 검출하기 위한 수단을 포함할 수 있으며 여기에 시약, 및 선택적으로 물질이 제공되어 임의의 1개 이상의 다음 단계를 수행한다: 분리, 고정화, 고유의 뉴클레오티드 서열 검출, 뉴클레오티드 서열의 연결, 및/또는 연결된 뉴클레오티드 서열을 검출하여 상기 동시-국소화 정보를 수득하는 단계. 상기 키트는 또한 본 발명의 방법을 수행하고 상기 키트를 이용하기 위한 설명서를 포함할 수 있다.

[0146] 본 발명의 키트는 단백질 디스플레이 라이브러리 (여기서 상기 라이브러리의 각 멤버는 고유의 뉴클레오티드 서열과 연관되어 있다)를 추가로 포함할 수 있다.

본 발명의 방법과 혼존 방법의 비교

[0148] 본 발명의 방법은 랜덤 샘플링과 리던더시가 높은 샘플링을 가능하게 하는, 시간 및 비용 효율적인 방식으로 상호반응체의 정보를 포착하는 새로운 접근법이다. 이는 심지어 크고, 다중-유니트 단백질 복합체의 정량적인 상호반응 데이터의, 동적이고, 원래의 세포 컨텍스트 기반의, 예를 들어, 생리학적, 천연 단백질-단백질 상호반응 기반의, 포괄적인 커버리지(coverage)를 제공한다. 이는 랜덤한 변수의 효과, 예컨대 비-특이적이고, 돌발적으로 상호반응하는 단백질의 검출을 억제한다. 이는 또한, 원래의 단백질-단백질 상호반응 외에 검출 원리에 포함되는 임의의 결합 반응 연관한 변수와 같은 변수의 효과를 없애준다. 본 방법은 베이트 단백질(bait protein)하고 상호반응할 뿐만 아니라, DNA, RNA 및 화학적 화합물과 상호반응하는 단백질의 시험관내 검출에 적합하다.

[0149] 라이게이션 프록시미티 어세이(ligation proximity assay, LGA)의 PCR 기반 검출법을 사용하여 표적의 상대적 발현을 측정한다. 단백질과 mRNA 발현 프로파일이 동일하지 않기 때문에, 관찰된 차이는 생물학적 과정면에서 볼 수 있다. 상기 검정법은 주어진 실험에서 수개의 표적을 검출할 수 있다. 하지만, 인공적으로 표지된 특이적 항체를 다수 제조하는 것이 필수적이기 때문에, 상호반응체 수준에서는 실행할 수 없다. 고도로 특이적인 항체 없이도, 교차-반응성은 특이적인 상호반응과 비-특이적인 상호반응을 명확하게 구별할 수 있고 라이게이션 프록시미티 어세이의 능력을 감소시킨다. 매우 큰 수의 표적의 평행 검출은 비용이 많이 들고 복잡하며 느리게 된다.

[0150] 다중복합화된 형태(multiplexed form)의 LGA가 개발되었다. 이 검정법에서는, 고체 지지체에 고정화된 항체가 포착 시약으로 작용하여 단백질의 복합체 혼합물로부터 항원을 국소적으로 강화시킨다. 세척 후, 한쌍의 프록시미티 라이게이션 어세이(proximity ligation assay, PLA) 프로브를 가한다. 이어서 추가로 세척하고 도입된 올리고뉴클레오티드를 근접하여 라이게이션 시킨다. 이는 3번의 결합 반응에 대한 필요성을 기본으로 더 높은 특이성을 가능하게 한다. PCR 증폭법을 사용함과 함께, 이는 높은 특이성과 감응성, 및 단백질 정량분석에 대한 광범위한 동적 범위를 가능하게 한다. 이 방법을 차세대 서열분석법(next generation sequencing, NGS)과 커플링시키면 단백질 존재비의 패턴이 디지털로 기록되어, 이를 사용하여 36개의 단백질 분석물의 동시 검출이 증명되었다.

[0151] LGA의 변형법은 미세소낭(microvesicle)과 같은 복합체 표적 구조물의 검출을 위한 극감응성이고 특이적인 검정법(4PLA)으로 기재되어 있으며 여기서는 표적이 먼저 고정화된 항체를 통하여 포착되고 후속해서 부착된 DNA 가닥이 있는 4개의 다른 항체를 사용함으로써 검출된다. 증폭가능한 리포터를 발생시키기 위하여 5개의 항체에 의한 우연한 결합의 요구조건으로 특이성과 감응성이 모두 증가된다.

[0152] 밀접한 근접(close proximity)을 사용하는 모든 타입의 프록시미티 라이게이션 기반 검정법은 매우 빈번한 입체적 구속으로 인하여 실험적 검증을 필요로 한다.

[0153] 본 발명의 방법의 경우, 낮은 특이성이 이슈가 되지 않으며, 검증 정보로도 사용될 수 있다. 본 발명의 방법은

동시-국소화(co-localization) 및 구획화(compartmentalization)를 이용하며, 여기서 증폭된 동일성 표지물은 구획으로 자유롭게 확산되어 더욱 이완된 입체 상태가 되도록 한다.

[0154] 이중 발현 재조합효소 기반 (Dual Expression Recombinase Based, DERB) 데스티니(destiny) 벡터는 개별적으로 관심의 대상인 단백질 오픈 리딩 프레임(open reading frame, ORF)을 삽입하기 위한 재조합효소 인식가능한 서열 2세트, 상호반응 검출을 위한 ORF가 있는 프레임의 프로모터와 리포터 태그 2세트를 코딩한다. 상기 벡터를 살아있는 세포 (원핵세포 및 진핵세포)에 도입시킴으로써 형광 공명 에너지 전달법(fluorescence resonance energy transfer, FRET) 또는 생분자 형광 보완법(bimolecular fluorescence complementation, BiFC)에 의해 단백질 상호반응의 검출이 가능하게 된다. DERB 플랫폼(DERB platform)은 재조합효소 기반 클로닝을 도입하고 원리증명(proof-of-principle) 실험을 통하여 검증된 벡터를 호환적으로 용인하여 미지의 상호반응을 식별함으로써 현재 상업화된 시스템보다 좋음을 보여준다. 상기 시스템은 많은 수의 선별된 상호반응을 요구하며, 결과적으로 많은 노력과 비용이 요구되어, 상호반응체 수준 및 인공 시험 조건 (융합 단백질 및 인공 프로모터)에서는 로봇 시스템에만 적합하다.

[0155] 이스트 투-하이브리드(yeast two-hybrid, Y2H) 선별법은 단백질-단편 보환 검정법, 또는 PCA의 특이적 실행법으로, 여기서는 단백질-단백질 상호반응의 식별이 각각 제3의 단백질 (예, DHFR, 이는 리포터로 작용한다)의 불완전한 단편에 공유결합에 의해 연결되어 있는 2개의 단백질 단편을 기본으로 한다. 상기 단백질 사이의 상호반응으로 리포터 단백질의 단편을 이들이 기능성 리포터 단백질 (이의 활성은 측정할 수 있음)을 형성하기에 충분히 밀접시킨다. 이 원리는 GAL4 전사 인자를 사용하는 이스트 투-하이브리드 선별법으로서, 수많은 다른 리포터 단백질에 적용시킬 수 있다. 상기 이스트 투-하이브리드 선별법은 효모의 핵 내부에서 인공 융합 단백질 간의 상호반응을 조사한다. 이 방법은 높은 거짓-양성 비율(false-positive rate)을 가지며, 이는 식별된 상호반응을 다른 수단으로 검증하는 것이 필수적이도록 한다. 이 방법은 세포 컨텍스트 범위를 넘어서며 천연 환경이 결여되어 있어 단백질의 세포 컨텍스트 특이적 변형에 의해 지배되는 상호반응 또는 친화성이 낮은 상호반응의 경우로 이의 용도가 제한된다. 상호반응체 수준에서는, 매우 큰 상호반응체 데이터세트를 만들 수 있도록 하기 위한, 추가적인 최적화 및 어레이를 사용할 필요가 있다. 이는 높은 비용을 포함하지만, 상기 검정법 탑입의 모든 제한점은 여전히 극복되지 않았다.

[0156] 오류 포지티브를 제거하기 위한 즉각적인 선택으로 라이브러리 스크린의 정확도를 개선한, 듀얼 베이트 시스템 (dual bait system)이 사용될 수 있다.

[0157] 꼬인-코일 매개된 혜테로다이머화 (coiled-coil mediated heterodimerization) 기능성 상호반응 트랩 검정법 (trap assay)이 기재된 바 있으며, 여기서는 꼬인-코일 혜테로다이머화 도메인이 모듈형 단백질 결합 도메인 대신 사용된다. 이는 기능적으로 연관있는 단백질-단백질 상호반응의 검증, 특이적 기질에 대해 효소를 지시, 및 기능적으로 중요한 상호반응 파트너에 대한 융합 라이브러리를 선별하는데 유용할 수 있다.

[0158] Y2H 선별법의 알려져 있는 단점에 대응하여, 포유동물 세포 기반 투 하이브리드(mammalian cell based two hybrid, M2H) 시스템이 개발되었다. 상기 M2H 시스템은 이트스 투-하이브리드의 것과 유사한데, 여기서는 상호반응이 관심의 대상인 각각의 단백질 쌍을 각각 DNA 결합 및 전사 트랜스-활성화 도메인에 융합시켜 조사한다. 포유동물 세포 기반 투-하이브리드 기법은 효모 기반 검정법과 비교하여, 수많은 장점을 가지며 알려져 있는 이슈 중 몇몇을 해결한다. 효모가 번역후 변형에 관여하는 중요한 단백질이 결여되어 있기 때문에, 이를 단백질에 기반한 상호반응은 측정할 수 없다. 또한, 세포 컨텍스트 특이적 상호반응체 데이터를 제공하기 위하여 수개의 상이한 포유동물 세포 컨텍스트를 사용할 수 있다. 그러나, 큰 데이터세트가 생산되기 어려워 상호반응체 수준의 상호반응은 너무 많은 수의 포유동물 세포 컬처 표본을 필수적으로 다뤄야하기 때문에 성취될 수 없다.

[0159] Y2H 선별법에서 단백질-단백질 상호반응 검출에서의 변화는 스티치-시퀀스(Stitch-seq)라 불리워지는, 특이적 PCR 기반 서열분석법을 사용하는 것이다. 이는 PCR 스티치법(PCR stitching)으로, 동일한 PCR 앰플리콘 상에 상호반응하는 단백질을 코딩하는 한 쌍의 서열을 놓는다. PCR 스티치법은 2라운드의 PCR로 이루어져 있다. 제1 라운드에서는, X와 Y (Y2H DB-X 및 AD-Y 벡터에 존재함)를 각각 DB- 및 AD 벡터-특이적 업스트림 프라이머로 증폭시킨다. 제1 라운드로부터의 앰플리콘을 주형으로 사용하여 82개-bp 링커 서열로 연결된 X 및 Y ORF로 구성된, 사슬같이 이어진 PCR 생성물(concatenated PCR product)을 생산한다. 상기 PCR 생성물을 모아 차세대 DNA 서열분석법으로 서열분석하여 스티치된 IST(stitched ISTs, sIST)를 생산한다. Stitch-seq가 일부 Y2H 프로토콜의 병목(bottleneck)을 제거하였지만, 여전히 중요한 단계와 연관된 문제점이 해결되지 않았다.

[0160] 상기 PCA에 대해 다음과 같은 개량: 1) 큰 라이브러리 ($>10^7$ 크기)의 분석을 용이하게 허용하며 이의 선택도가 용이하게 "조율되고(tuned)", 개량되며/되거나 모니터될 수 있는 리포터 유전자 (및 이들의 발현을 검출하는 방

법), 2) 다중 상호반응의 동시적이고 독립적인 측정 방법 (상이한 리포터 유전자의 발현에 의해 판단됨에 따라), 및 3) a) 라이브러리 대 라이브러리 실험을 수행하기 위한 효율적이고 자동화 가능한 방법 및 b) 원핵세포 PCA에서 수행된 임의의 선별/선택으로부터 포지티브 후보물질의 분석을 단순화하기 위한 방법을 제공하는, 파아지미드(phagemid)-기반 시스템을 사용하여 라이브러리의 구축을 기재하고 있는 특허. 라이브러리 대 라이브러리를 선별하기 위한 상기 파아지미드-기반 기술을 사용하게 되면 교차 라이브러리를 포함하게 되며, 예를 들면 세포의 베이트 라이브러리를 파아지의 먹이 라이브러리로 (각 세포가 평균적으로 1개의 파아지에 의해서만 감염되도록 파아지보다 과량의 세포를 사용하여) 감염시키고 리포터 유전자의 활성화된 발현을 찾는다. 이는 큰 규모의 상호반응체 스캔에 대해 획기적인 단계이지만, 동적이거나 세포 컨텍스트를 기반으로 하지 않는다.

[0161] 효모에서 수행되는 라이브러리 대 라이브러리 실험의 경우, 출발 반수체 세포 둘 다로부터의 DNA가 들어있는 2 배체 α 세포를 형성시키는 단계가 사용된다. 따라서, 먹이 하이브리드의 라이브러리가 들어있는 세포는 시험 베이트 하이브리드가 들어있는 α 세포와 짹을 이를 수 있다. 이는 형질전환 효율 이슈를 제거하지만, 이상적인 상호반응체 스캔의 다른 요구조건은 해결하지 못한다.

[0162] 소수성 코어에 인접한 위치에서 반-랜덤화된, 2개의 류신 지퍼 라이브러리를 효소 쥐의 디하이드로폴레이트 환원효소 (mDHFR)의 2개의 디자인된 단편 중 하나에 유전적으로 융합시키고 이. 콜리(E. coli) 중으로 동시에 형질전환시켰다. 상기 라이브러리 폴리펩티드 사이에서의 상호반응은 mDHFR의 효소 활성의 재구성에 요구되며, 이는 세균이 성장하도록 한다. 하지만, 이 전략은 세균 세포에서 성취될 수 있는 형질전환 효율에 의해 제한된다.

[0163] GST-풀다운(pulldown) 검정법을 96개-웰 필터 플레이트 포맷(format)에 적응시키는 것이 또한 고안되었다. 다중-웰 필터 플레이트를 사용하게 되면 전통적인 단일 투브 검정법에서 하는 것보다 시약을 덜 사용하고 더욱 효율적인 샘플 가공법을 사용하여 훨씬 더 짧은 시간에 더 많은 샘플을 검정할 수 있도록 한다. 이런 타입의 검정법은 기술적 병목으로 야기되는 문제점의 일부를 해결하지만; 상기 시스템을 사용해서는 요구되는 매우 큰 데이터세트를 생성시킬 수 없다.

[0164] 텐덤 친화성 정제(Tandem affinity purification, TAP) 방법은 동시-면역침전법(co-immunoprecipitation)의 더욱 특이적인 버전으로 알려질 수 있으며, 단백질 상호반응의 고효율 식별을 가능하게 한다. 상기 방법의 정확도는 작은-규모 동시-면역침전 실험의 정확도에 버금갈 수 있으며 정확한 세포 컨텍스트 내에서 상호반응이 검출된다. 하지만, 이는 2개의 연속적인 단백질 정제 단계를 필요로 하기 때문에, 일시적인 단백질-단백질 상호반응을 용이하게 검출할 수 없다. TAP 방법은 TAP 태그를 연구 대상인 단백질의 C-말단에 융합시키는 것을 사용한다. 상기 TAP 태그는 N-말단으로부터의 칼모듈린 결합 웨პ티드(calmodulin binding peptide, CBP), 이어서 담배 식각 바이러스 프로테아제(tobacco etch virus protease, TEV 프로테아제) 절단 부위 및 프로테인 A로 이루어져 있다. 이는 정확한 세포 컨텍스트에서 정량적으로 단백질 파트너를 실제 측정할 수 있도록 하지만, 상호반응체 수준에서 이 방법은 많은 노력을 필요로 하며 사용되는 작제물로 전체 프로테옴(proteome)을 커버하기 위해 많은 비용을 발생시킨다.

[0165] 이런 이슈에 대응하여, 고성능 멀티베드 이온-교환 크로마토그래피, 슈크로오스 구배 농축법 및 등전점 전기영동법을 포함한, 다중 분리 기법을 사용하여 가용성인 인간 단백질 상호반응체의 체계적이고, 고도로 집중적인 생화학적 분획화를 수행하기 위한, 태그가 없는 (tagless) 전략이 개발되었다. 이 방법은 신뢰할만한 데이터세트를 생산하기 위한 입증 및 통계적 분석을 필요로 하며, 또한 현저한 양의 샘플 물질을 필요로 하고 이의 비용은 랜덤 샘플링과 높은 리던던시의 샘플링을 위한 이의 사용을 제한한다.

[0166] 단일 실험으로 수천 개의 단백질 간의 새로운 상호반응을 신속하게 식별할 수 있는 능력을 제공하는, 단백질 상호반응을 식별하고 특징화하기 위한 새로운 접근법으로 단백질 마이크로어레이가 도입되었다. 어레이 상에 각 단백질의 위치와 동일성이 알려져 있기 때문에, 상호반응 지도(map)가 단백질 어레이의 반복적인 조사 (iterative probing)로부터 신속하게 개발될 수 있다. 단백질 마이크로어레이 실험은 하루 내에 수행되고, 상호반응이 수천개의 다른 단백질의 컨텍스트로 평가되기 때문에, 마이크로어레이 상의 상호반응 프로파일링은 신규 단백질 상호반응이 밝혀지는 속도를 크게 가속시킬 수 있다. 추가로, 단백질 마이크로어레이 실험의 시험관내 특성은 단백질 농도, 번역후 변형(post-translational modification), 및 보조인자의 존재와 같이 단백질 상호반응에 영향을 주는 조사 조건에 대한 제어를 허용하는데, 이는 이스트 투-하이브리드 선별법과 같은 다른 방법으로 가능할 수 없다. 하지만, 시간 접근에서 전통적인 하나의 프로브는 큰 상호반응체 수준의 실험에 적합하지 않다.

[0167] 단백질 어레이 기반 검출 버전이 사용되는데, 여기서는 세포의 단백질 용해물 또는 합성 웨პ티드 믹스를 고정화된 베이트 단백질/웨პ티드가 있는 단백질 어레이에 인가한다. 다양하고 엄격한 조건하에서 비특이적인 단백질/웨

티드를 세척하여 버리고 상기 베이트 단백질/펩티드와 특이적으로 상호반응하는 단백질만 칩에 남긴다. 마지막으로, 상기 포착된 상호반응하는 단백질/펩티드 복합체를 SELDI-TOF 질량 분석법으로 분석하고 이들의 동일성은 이들의 예측되는 독특한 질량으로 식별한다. 이는 매우 기대가 되는 접근법이지만, SELDI-TOF에 의해 단백질 서열분석이 여러가지 인자 (단백질의 양, 분리, 번역후 변형)에 의해 제한되며 상기 상호반응은 천연 컨텍스트가 결여되어 있다.

[0168] 단백질 혼합물을 또한 고체 지지체에 고정화시키고, 복수의 표지되지 않은 단백질-단백질 상호반응 도메인과 적절한 결합 조건하에서 접촉시킬 수 있다. 적어도 1개의 표지되고 선택된 단백질-단백질 상호반응 도메인의 존재 하에서, (상기 표지된 단백질-단백질 상호반응 도메인은 표지되지 않은 단백질-단백질 상호반응 도메인과 다르다) 표지된 단백질-단백질 상호반응 도메인의 결합을 측정한다. 이 방법은 상호반응 도메인에 특이적이며, 이는 이의 적용을 제한한다.

[0169] 차세대 서열분석법(NGS)과 조합된 세포-유리식 (cell-free) 디스플레이 기술은 상호반응체 데이터세트의 커버리지와 신뢰도를 둘 다 향상시킬 수 있다. 완전한 세포-유리식 방법은 고효율이고 큰 검출 공간을 제공하며, 클론을 사용하지 않고도 상호반응을 시험할 수 있다. NGS에 의해 제공되는 정량적 정보는 오류 포지티브의 수를 감소시킨다. 이 방법은 베이트 단백질뿐만 아니라, DNA, RNA 및 화학적 화합물과 상호반응하는 단백질의 시험관내 검출에 적합하다. 이 방법은 cDNA 라이브러리 (세포 및 조직으로부터 추출된 것)로 완벽하게 시험관내 처리하는 단계 및 표적 단백질에 대해 선택하여 NGS용 선택된 cDNA 서열을 얻는 단계를 이용한다. 이 방법을 사용하는 선택은 세포-유리 조건하에서 수행되며, 후속되는 NGS에 의한 서열분석은 임의의 종류의 세포를 사용하는 클로닝 단계에 제한되지 않는다. 이 방법은 세포 컨텍스트를 넘어 시간 접근법으로 1개를 사용하며, 이는 상호반응체 수준에서 큰 데이터세트를 생성할 수 있는 이의 능력을 제한한다.

[0170] 다른 세포 유리 검정법은 소위 인테인(intein) (이는 시험관내에서의 단백질 트랜스-스플라이싱을 지시할 수 있는 펩티드 서열이다)을 이용한다. 인테인은 단백질 스플라이싱 중 단백질 전구체로부터 절제된 단백질 전구체 중의 매개 단백질 서열(intervening protein sequence)이다. 2개의 하이브리드 융합 작제물(Two hybrid fusion construct)이 제공되며, 여기서 하나는 제1 시험체와 N-말단 인테인 단편 또는 N-인테인을 가지며, 다른 것은 제2 시험체와 C-말단 인테인 단편 또는 C-인테인을 갖는다. 또한, 융합 작제물 중 하나 또는 둘 다 융합 작제물의 트랜스-스플라이싱 때 검출가능한 변화를 일으키는 리포터를 가질 수 있다. 처리율 및 세포 컨텍스트 유리 특성 둘 다 이 방법의 확실한 단점이다.

[0171] 단백질-DNA 또는 단백질-단백질 상호반응을 연구하는데 사용되는 다른 방법으로 파아지 디스플레이(phage display) 방법이 있다. 디스플레이된 단백질의 DNA를 코딩하는 실모양의 박테리오파아지 (예, M13)의 표면에 단백질이 디스플레이된다. 표적 단백질 또는 관심의 대상인 DNA 서열을 고체 지지체에 고정화시키고 이를 사용하여 표적에 결합하는 후보물질에 대한 파아지-디스플레이된 단백질의 라이브러리를 친화성-강화(affinity-enrich)시킨다. 이 방법은 단백질-DNA와 단백질-단백질 상호반응을 모두 식별하고 특성화하는데 사용되었다. 파아지 디스플레이는 단백질-단백질 상호반응 데이터를 추론하기 위하여 다중 사이클을 필요로 하는 강화 과정(enrich process)이다. 상기 강화 과정은 시험관내에서 수행되며, 이는 상호반응을 편향시키고 검출에 있어서 친화성이 높은 상호반응에 유리하다. 특정 단백질 (특히 더 큰 것)은 파아지 디스플레이에 의한 분석에 잘 맞지 않는다. 세균-기반 OFR 파아지 디스플레이의 주요 단점은 파아지 표면에 디스플레이된 단백질이 적절한 번역후 변형, 예컨대, 글리코실화가 결여되어 있다는 것이다.

[0172] 당해 분야에서 사용되는 바와 같은, 파아지 디스플레이는 상기 기재된 바와 같이, 천연 결합 반응을 재현하는 것으로 제한된다. 하지만, 파아지 디스플레이로 발현된 항체 레퍼토리는 치료용 항체에 대한 약리학적 선두물질을 제공하거나 진단 검정용 항체를 검출하는데 있어서 극히 성공적이다. 본 발명은 후자의 컨텍스트로 파아지 디스플레이를 이용하며, 여기서 매우 큰 파아지 항체 라이브러리는 거의 모든 식별가능한 표적을 커버하는 복수의 특이성을 포함한다.

[0173] 파아지 디스플레이의 변화는 상호반응 파트너의 선택 및 식별 방법에 관한 것이다. 표적 분자 (리간드)를 고체 상 담체의 표면에 고정화시켜 이들이 2차원 등급으로 위치에 정착할 수 있도록 하고 단백질 디스플레이 바이러스와 접촉시킨다. 상호반응 파트너는 고정화된 리간드와 상호반응 파트너 간의 결합 위치를 검출하고 측정함으로써 식별된다. 기재된 바람직한 검출 방법은 표면 플라즈몬 공명(surface plasmon resonance: SPR)이다.

[0174] 상호반응하는 단백질을 단리/식별하기 전에 단백질 산호반응을 자리에 "고정(fix)"시키기 위하여 흔히 화학적 가교결합이 사용된다. 상기 적용에 일상적인 가교-링커(cross-linker)에는 비-절단성 [NHS-에스테르] 가교 링커, [비스-суль포숙시니미딜 수베레이트(bis-sulfosuccinimidyl suberate)] (BS3); BS3의 절단성 버전, [디티

오비스(술포숙시니미딜 프로피오네이트)(dithiobis(sulfosuccinimidyl propionate)] (DTSSP); 및 ChIP 검정법에서 상호반응을 고정시키는데 인기있는 [이미도에스테르(imidoester)]가교링커 [디메틸디티오비스프로피온이미데이트(dimethyl dithiobispropionimidate)] (DTBP)가 포함된다.

[0175] 표면 고정화된 이중사슬 DNA에 결합하며, 관심의 대상이 되는 DNA 모티브(motif) 서열을 함유하는, 계놈적으로 코딩된 펩티드의 파아지-디스플레이 라이브러리를 사용하여 예측된 전사 조절 인자에 특이적으로 결합하는 단백질을 식별하는 기술이 개발되었다. 특이적인 DNA-단백질 상호반응에 대해 강화(enrichment)시킨 후, 결합된 파아지를 증폭시키고, 상기 강화된 파아지로부터의 삽입물을 서열분석하여 표지화 및 DNA 마이크로어레이로의 하이브리드화를 이용하여 상호반응하는 단백질을 결정한다.

[0176] chemFET 어레이를 사용하여 1개 이상의 분석물을 전하는 바에 따라 측정하였다. 상기 어레이는 항원으로의 항체의 결합을 포함하여 관심의 대상인 화학 공정 또는 화학 공정들에 대해 연관있는 정보를 제공하는 임의의 다양한 화학 물질을 포함할 수 있다. 일부 양태로, 단지 분석물의 존재를 검출하는 것 외에, 1개 이상의 분석물의 수준 또는 농도를 측정할 수 있는 능력은 화학적 과정 또는 과정들과 연관하여 가치 있는 정보를 제공한다.

[0177] 상기 설명으로부터 자명한 바와 같이, 본 발명은 단백질-단백질 상호반응의 검출 및 특징화와, 단백질-단백질 상호반응을 조절할 수 있는 화합물의 선택을 위한, 강력하고, 다목적의, 시험관내 시스템을 제공한다. 상기 시스템은 매우 편리하게 사용할 수 있으며 고효율의 선별 과정에 용이하게 적응시킬 수 있다.

[0178] 본 명세서에 언급된 모든 공보 및 특허출원은 본 발명이 속해있는 기술 분야의 숙련가의 수준의 지표이다. 모든 공보와 특허출원은 본 발명에서 각각의 개별 공보 또는 특허출원이 참고문헌으로 포함된다고 구체적이고 개별적으로 디스플레이되는 것과 동일한 정도로 참고문헌으로 포함된다.

[0179] 전술한 발명이 일부 상세하게 명확한 이해를 목적으로 설명의 방식으로 기재되었지만, 특정 변화와 변형은 첨부되는 특허청구의 범위의 범주 내에서 실행될 수 있음이 분명하다.

[0180] 실시예 1

[0181] 파아지 용해와 액적 디지털 PCR에서의 검출(Phage lysis and detection in droplet digital PCR)

[0182] pBluescript II SK(+) 파아지미드 벡터 (Agilent, 212205) f1 기원(origin) (+) 배향(orientation), 숙주 균주에서 Sac-->Kpn 폴리링커 배향(polylinker orientation): XL1-Blue MRF'를 사용하여 pBluescript II SK(+) 파아지를 생성시키고 M13K07 헬퍼 파아지 (NEB, N0315S)는 구입하였다. 상기 파아지들을 적정하여 감염성 파아지의 수를 측정하였다. 상기 파아지들을 10의 인수(factor of 10)로 순차적으로 희석시켜 구획 희석 당 단일 파아지보다 더 낮게 만들었다. 상기 희석액을 QX200 액적 디지털 PCR (ddPCR™) 시스템을 사용하여 디지털 PCR을 수행하였다. 간략하면, 프로브 5'CTGGTAGCGGTGGTTTT3' (pBluescript II SK(+) 특이적인 프로브 FAM-MGB 표지됨), 5'CCGTCAATTTACCTCC3' (M13K07 특이적인 프로브 VIC-MGB 표지됨)가 있는 5'CTCAAGTCGGTGACGGTGAT3' (M13K07 특이적인 정방향), 5'GACAAAAGGGCGACATTCAA3' (M13K07 특이적인 역방향) 및/또는 5'TCTTGATCCGGCAAACAAAC3' (pBluescript II SK(+) 특이적인 정방향), 5'TTTCTGCGCGTAATCTGCT3' (pBluescript II SK(+) 특이적인 역방향)를 사용하여 상기 구획화된 파아지를 증폭시키고, 상기 증폭을 2개의 상이한 채널에 기록하며 액적 생성, PCR 및 검출을 제조업자의 프로토콜에 따랐다.

[0183] 효과적인 파아지 용해와 카운트의 단일 파아지 검출 포아송 분포(single phage detection Poisson distribution of counts)가 검출되며 이는 단일 파아지 검출 감응성을 나타낸다.

[0184] 실시예 2

[0185] 다이머화(dimerisation) PCR

[0186] pBluescript II SK(+) 파아지미드 벡터 (Agilent, 212205) f1 기원 (+) 배향, 숙주 균주: XL1-Blue MRF'에서 Sac-->Kpn 폴리링커 배향을 사용하여 pBluescript II SK(+) 파아지를 생성시키고 M13K07 헬퍼 파아지 (NEB, N0315S)는 구입하였다. EZ-Link Sulfo-NHS-Biotin (Thermo, 21326)을 사용하여 제조업자 지침에 따라서 20몰 과량의 비오틴으로 파아지들을 표지시키고 문헌: Dong D, Sutaria S, Hwangbo JY, Chen P. A simple and rapid method to isolate purer M13 phage by isoelectric precipitation. Appl Microbiol Biotechnol. 2013 Sep;97(18):8023-9에 기재된 바와 같이 등전점 침전법을 사용하여 침전시켰다. 동수의 정제된 파아지를 혼합하여 10e+6/파아지mL(10e+6/ml of phages)의 농도를 수득하여 몰 당량으로 아비딘(A9275-1MG, Sigma-Aldrich)과 합하거나 혼합물로 사용하여 디지털 PCR (QX200 액적 디지털 PCR (ddPCR™) 시스템)을 수행하였다. 간략하면, 프로브 5'CTGGTAGCGGTGGTTTT3' (pBluescript II SK(+) 특이적인 프로브 FAM-MGB 표지됨),

5'CCGTCAATATTACCTTCCC3' (M13K07 특이적인 프로브 VIC-MGB 표지됨)가 있는 5'TAACGTGGGAATGGTGCTTCAGTCGGTGACGGTGAT3' (M13K07 특이적인 정방향), 5'GACAAAAGGGCGACATTCAA3' (M13K07 특이적인 역방향) 및 5'GAAGCACCATTCCACGTTATCTGATCCGCAAACAA3' (pBluescript II SK(+) 특이적인 정방향), 5'TTTCTGCGCGTAATCTGCT3' (pBluescript II SK(+) 특이적인 역방향)를 사용하여 상기 구획화된 파아지를 증폭시키고, 상기 증폭을 2개의 상이한 채널에 기록하며 액적 생성, PCR 및 검출을 제조업자의 프로토콜에 따랐다. 증가된 연결(linkage) 개수(count)는 비오틴-아비딘 결합으로 인하여 아비딘의 존재에서 예측되었다. 다이머화된 PCR 생성물을 검출하기 위하여 상기 증폭된 DNA를 제조업자의 권장에 따라서 추출하고 실시간 PCR 장치에서 프라이머 5'TTTCTGCGCGTAATCTGCT3' (pBluescript II SK(+) 특이적인 역방향), 5'GACAAAAGGGCGACATTCAA3' (M13K07 특이적인 역방향)과 프로브 5'CTGGTAGCGGTGGTTT3' (pBluescript II SK(+) 특이적인 프로브 FAM-MGB 표지됨), 5'CCGTCAATATTACCTTCCC3' (M13K07 특이적인 프로브 VIC-MGB 표지됨)를 사용하여 PCR을 수행하였다. 정확하게 다이머화된 생성물만 증폭될 수 있으며 예측되는 증폭 시그널이 두 프로브에 의해 생성된다.

[0187] 실시예 3

[0188] 단백질/단백질 복합체의 구획 기반 식별(Compartment based identification of proteins/protein complexes)

[0189] (a) 표준 방법에 의해 단백질 복합체의 추출을 수행한다 - 임의의 분리 방법 (이는 부분적으로 보존된 단백질 복합체라도 제공한다) 및 단백질 복합체를 부분적으로 정제하는 방법 (예, 이들의 번역후 변형 등에 따라서)을 포함함.

[0190] (b) 항체 라이브러리 파아지를 단백질 복합체와 합하고 희석시켜 구획화하고 비결합 파아지에 대해 단일 파아지 수준의 분리를 달성한다. 상기 구획화 (예, 다수의 반응의 효과적인 분리)는 단일 비결합 파아지의 포아송 분포 기반 분리에 기반하며; 예를 들어, 에멀젼 및 마이크로어레이가 현재 알려진 기술 상태에서 최선의 방법이다. 유일한 필요조건은 파아지와 단백질 복합체를 사용한 최소의 비특이적 동시-국소화이다. 뉴클레오티드 서열은 파아지의 효과적인 열 용균(heat lysis)과 PCR 과정을 연결하여 연결된다.

[0191] (c) 구획 당 파아지 (결합 및 비결합)의 항체 코딩 (파아지 라이브러리에서 항체 유전자의 임의의 CDR 영역에 기반, 바람직하게는 CDR3 기반), 특이적 DNA 단편을 증폭시키고 (일반적인 프라이머로 가능한 다이머화를 사용) 연결한다 (다이머화 PCR은 예를 들면, 실시예 2와 4를 참조한다).

[0192] (d) 상기 생성된 다이머화된, 심지어 정제되지 않은 PCR 앰플리콘을 바람직하게는 마이크로어레이 공동(cavity) 또는 기타 수단으로부터 반응물을 에멀젼 또는 물리적 제거로부터 DNA를 추출하는 방법을 사용하여 합한다.

[0193] (e) 다수(multitude)의 PCR 다이머 앰플리콘의 연결 정보가 고도로 평행이고 정량적인 방식으로, 바람직하게는 차세대 DNA 서열분석에 의해 밝혀진다.

[0194] (f) 각 파아지의 예정된 결합 특성 정보 (실시예 5 및 7 참조)와 결합된 파아지의 연결 정보를 사용하여 배경을 뺀 연결 정보 (난잡하거나 랜덤하거나 돌발적인 상호반응 제거)를 기본으로 유의적인 상호반응의 측정; 이들의 공지된 해리 상수 정보에 의해 가능하게 칭량된 동일한 결합 특성 정보를 사용하여 상이한 파아지의 과잉의 연결 정보를 기본으로 상호반응의 식별 및 여과; 식별되고 여과된, 가능하게 칭량된 연결 정보를 기본으로 다이머 및 멀티머 상호반응의 측정; 정량적인 연결 정보를 기본으로 단백질과 단백질 복합체의 상대적 존재비의 측정; 동일한 결합 특성 정보를 사용하여 수개의 파아지의 검출의 과잉 측정치를 기본으로 다이머와 멀티머 상호반응 및 단백질과 단백질 복합체의 상대적 존재비의 측정의 통계적 오류의 식별, 계산을 포함한, 통계적 기준으로 상호반응체를 컴퓨팅한다.

[0195] 실시예 4

[0196] 대조 단백질/단백질 복합체의 구획 기반 식별(Compartment based identification of control proteins/protein complexes)

[0197] 파아지미드 항체 라이브러리 (~3x10⁹) (Dudgeon K, Famm K, Christ D. Sequence determinants of protein aggregation in human VH domains. Protein Eng Des Sel. 2008 Oct 28.) KM13 헬퍼 파아지, TG1Tr 세균 균주 및 항-베타-갈락토시다제 & 항-소의 유비퀴틴 항체(anti-bovine ubiquitin antibodies) 디스플레이 파아지를 포함하여, M13 파아지를 디스플레이하는 항체를 Source BioScience (6001_hDAb)로부터 구입하였다. 대조용 파아지를 NGS (ThermoFisher Scientific PGM, Ion Xpress™ Plus Fragment Library Kit, 4471269)로 서열분석하고 이들의 결합 친화성은 각 항원에 대해 ELISA로 입증되었다. 항원은 LifeSensor (#S1280)로부터의 모노-비오

티닐화된 유비퀴틴 (b-UBI), 베타-갈락토시다제 비오틴 표지된 (G5025)(b-BGAL)이다. 계란 흰자로부터의 아비딘 (A9275)도 Sigma-Aldrich로부터 구입한다.

[0198] 서열분석 결과를 기준으로 일반적인 프라이머가 디자인된다: 일반적인 정방향 - CCAAGAACACGCTGTATCTGCA; 다이머화할 수 있는 일반적인 역방향 프라이머 - TGGCATCCATTGTAGAGGTGAGACGGTGACCAGGGTTCC 및 ACCTCTACAATGGATGCGCAGAGACGGTGACCAGGGTTCC. 다이머화된 생성물을 검출하기 위하여 다이머 특이적인 실시간 PCR 반응이 디자인된다: 정방향 - AGTTGGAGTCTTGGGTCAGG, 역방향 - AGGTGGGTCGATGTTGACTACTG 및 프로브 - FAM TCTCACCTACAATGGAT MGB.

[0199] 대조용 파아지 특이적 프로브가 또한 디자인된다: 항-베타-갈락토시다제 - FAM GCTAGGGCTATGTATCC MGB; 항-소의 유비퀴틴 - VIC TGGGTCGATGTTGACTAC MGB.

[0200] 지침에 따라서 대조용 파아지를 증폭시킨다 (항-베타-갈락토시다제 $3.8 \times 10^{12}/\text{ml} = 6.48 \text{ nM}$, 항-소의 유비퀴틴 $4.0 \times 10^{12}/\text{ml} = 6.7 \text{ nM}$). 아비딘 (36 nM) (또는 제외시켰음), b-UBI (72nM) 및 b-BGAL (모노머의 경우 72 nM) 을 함하고 1시간 동안 실온에서 배양시켜 복합체를 형성시켰다. 10배 희석시킨 복합체를 합하여 밤새 1.5 nM의 대조용 파아지와 함께 배양시켰다. 파아지 결합된 복합체를 2×10^6 배로 희석시키고 (QX200 Droplet Digital PCR (ddPCR™) System)의 프로토콜에 따라서 애멸전 액적을 발생시키고 PCR 조건을 사용하여 증폭시킨다: 프로브 (no dUTP) (186-3023)에 대한 ddPCR Supermix, 일반적인 정방향 프라이머 농도는 800 nM이고, 다이머화할 수 있는 일반적인 역방향 프라이머 농도는 50 nM이다. 일부 경우에 있어서, 대조용 파아지 특이적 프로브를 또한 250 nM로 첨가한다. 상기 증폭된 액적을 제조업자의 프로토콜에 따라서 클로로포름 추출하여 증폭된 다이머 생성물을 회수한다.

[0201] 상기 다이머는 다이머 특이적 실시간 PCR에 의해 성공적으로 검출되어 이들의 정확한 다이머화된 구조가 증명된다. 대조용 특이적 프로브가 애멸전 PCR 반응에 포함되었을 경우 대조용 파아지와 아비딘/항원 복합체 결합으로 인하여 아비딘의 존재하에서 증가된 '연결(linkage)'-들이 검출되며, 이는 항-베타-갈락토시다제와 항-소의 유비퀴틴 파아지의 검출이 동일한 액적에서 우연한 경우에서보다 더 높은 비율로 국소화된 것으로 나타났다.

[0202] 실시예 5

[0203] 항체 파아지 라이브러리의 결합 특성 정보의 예비측정(Predetermining the binding characteristics information of an antibody phage library)

[0204] (a) 항체 (파아지 라이브러리, 이는 측정할 결합 특성 정보)와 cDNA 라이브러리 파아지 (cDNA로 구성, 이는 항체 파아지 라이브러리의 결합 특성 정보에 대해 측정될 것임)를 합하여 항-cDNA 파아지 복합체를 희석시키고 구획화하여 비결합 파아지 - 추가의 세부사항은 실시예 3의 (b) 부분을 참조 - 의 단일 파아지 분리 수준을 성취한다.

[0205] (b) 구획 당 파아지(결합 및 비결합)의 항체 코딩 (파아지 라이브러리의 항체 유전자의 임의의 CDR 영역을 기반으로, 바람직하게는 CDR3 기반) 및 cDNA 코딩 (cDNA 단편) DNA 서열을 증폭시키고 (다이머화할 수 있는 일반적인 프라이머를 사용) 연결한다 (예로서 다이머화 PCR, 실시예 2와 4를 참조).

[0206] (c) 상기 생성된 다이머화된 PCR 앰플리콘을 바람직하게는 마이크로어레이 공동으로부터 반응물을 애멸전 또는 물리적 제거로부터 DNA를 추출하는 방법을 사용하여 합한다.

[0207] (d) 다수의 PCR 다이머 앰플리콘의 연결 정보는 고도로 평행이고 정량적인 방식으로, 차세대 DNA 서열분석에 의해 밝혀진다.

[0208] (e) 결합된 파아지의 연결 정보를 사용하여 배경을 뺀 연결 정보 (난잡하거나 랜덤하거나 돌발적인 상호반응 제거)를 기본으로 유의적인 상호반응의 측정; 통계적으로 유의적인 상호반응(significant interaction)을 기본으로 유의적인 항체(significant antibody)-cDNA 결합의 식별; 검출된 cDNA 단편, 추론되어 검출된 단백질을 포함하여 각각의 검출된 항체-파아지에 대한 결합 특성의 측정; 동일한 결합 특성 정보를 사용하여 수 개의 파아지의 검출의 과정 측정치를 기본으로 결합 특성 정보의 통계적 오류의 식별, 계산을 포함하는 cDNA-파아지 라이브러리에 대한 항체 파아지 라이브러리의 결합 특성 정보를 컴퓨팅한다.

[0209] 실시예 6

[0210] 항체 파아지 라이브러리의 강화(Enrichment of an antibody phage library)

- [0211] (a) cDNA 라이브러리 파아지 (강화된 항체 파아지 라이브러리로 검출할 필요가 있는, cDNA로 구성되어 있음)를 분리가능한 수단(비결합 항체 파아지의 분리를 위하여)으로, 바람직하게는 마이크로비드 상에, 고정화시킨다.
- [0212] (b) 항체 라이브러리 파아지 (강화시키고자 하는, 파아지 라이브러리)를 고정화된 cDNA-파아지와 합하여 결합 및 비결합 파아지를 분리시킨다.
- [0213] (c) 비결합 항체 파아지를 바람직하게는 세척에 의해 제거한다.
- [0214] (d) 결합된 항체 파아지를 용출시키고 선택적으로 적합한 수단으로 증폭시켜 역가가 높은 제제(hight titer preparation)를 수득한다.
- [0215] (e) 선택적으로, 강화된 항체 파아지의 역가가 높은 제제를 다음 라운드의 강화처리한다(the high titer preparation of enriched antibody phages are subjected to next round of enrichment).
- [0216] (f) 선택적으로, 상기 용출된 결합된 파아지를 실시예 5에 기재된 방법을 사용하여 cDNA 파아지에 대해 입증한다.
- [0217] 실시예 7
- [0218] 강화된 항체 파아지 라이브러리의 결합 특성 정보의 예비측정(Predetermining the binding characteristics information of an enriched antibody phage library)
- [0219] 파아지미드 항체 라이브러리 (~3x10⁹) (Dudgeon K, Famm K, Christ D. Sequence determinants of protein aggregation in human VH domains. Protein Eng Des Sel. 2008 Oct 28.) KM13 헬퍼 파아지 및 TG1Tr 세균 균주를 포함하여, M13 파아지를 디스플레이하는 항체를 Source BioScience (6001_hDAb)로부터 구입하였다. 상기 라이브러리를 증폭시키고, KM13 헬퍼 파아지로 감염시켜 상기 파아지를 프로토콜 (Lee CM, Iorno N, Sierro F, Christ D. Selection of human antibody fragments by phage display. Nat Protoc. 2007;2(11):3001-8.)에 따라서 하베스트하였다. PhD12 파아지 디스플레이 웹티드 라이브러리 (E8110S)와 E. coli ER2738 숙주 균주는 New England Biolabs로부터 구입하였다. PhD12 라이브러리를 LB/IPTG/Xgal 플레이트에 플라크가 형성되도록 하고 50 개의 플라크를 집어내서 합하였다 (항원 베이트 라이브러리). 항원 베이트 라이브러리를 미세역가판에 흡착시키고 전체 Source BioScience (6001_hDAb) 라이브러리를 사용하여 Source BioScience (6001_hDAb) 라이브러리 프로토콜에 따라서 패닝(panning)을 수행하였다. 모두 612개의 클론을 LB/암피실린 플레이트에 플레이팅시키는데, 이를 증폭시키고, KM13 헬퍼 파아지로 감염시켜 하베스트하였다 (강화된 항체 라이브러리).
- [0220] 상기 파아지의 서열을 기본으로 일반적인 프라이머가 디자인된다: 일반적인 정방향 Source BioScience (6001_hDAb) 라이브러리 특이적 - CCAAGAACACGCTGTATCTGCA; 다이머화할 수 있는 일반적인 역방향 프라이머 Source BioScience (6001_hDAb) 라이브러리 특이적 - TGCGCATCCATTGTAGAGGTGAGACGGTACCGAGGTCC- 및 일반적인 정방향 PhD12 파아지 디스플레이 웹티드 라이브러리 특이적 - CGCAATTCCCTTAGTGGTACCTTT; 다이머화할 수 있는 일반적인 역방향 프라이머 PhD12 파아지 디스플레이 웹티드 라이브러리 특이적 - ACCTCTACAATGGATGCGCATCTGTATGGGATTTGCTAACAACT.
- [0221] 다이머화된 생성물을 검출하기 위하여 다이머 특이적 실시간 PCR 반응이 디자인된다: 정방향 - CGGACTGTTGAAAGTTGTTAGCA, 역방향 - GGTCAACCGTCTCACCTCTAC 및 프로브 - VIC-CATACAGATGCGCATCC-MGB.
- [0222] 10¹² 개의 항원 베이트 라이브러리와 강화된 항체 라이브러리 파아지를 합하여 실온에서 밤새 배양시켰다. 파아지 복합체를 2×10⁶ 배로 희석시키고 (QX200 Droplet Digital PCR (ddPCR™) System)의 프로토콜에 따라서 애벌전 액체를 발생시키고 PCR 조건을 사용하여 증폭시킨다: 프로브 (no dUTP) (186-3023)에 대한 ddPCR Supermix, 일반적인 정방향 프라이머 농도는 800 nM이고, 다이머화할 수 있는 일반적인 역방향 프라이머 농도는 50 nM이다. 상기 증폭된 액체를 제조업자의 프로토콜에 따라서 클로로포름 추출하여 증폭된 다이머 생성물을 회수한다. 상기 증폭된 다이머 생성물을 NGS 서열분석하고 특이적인 베이트 라이브러리와 강화된 항체 라이브러리 다이머화된 생성물이 검출되는데 이는 항원 베이트 라이브러리의 멤버 간의 상호반응을 기본으로 특이적 서열화되고 항체 라이브러리가 헤지(hedge)되었음을 나타낸다.
- [0223] 실시예 8
- [0224] 항체 파아지 라이브러리의 멤버의 정량적 결합 정보의 측정(Determining the quantitative binding information of the members of an antibody phage library)

- [0225] (a) 단계 b에서, 수 개의 정량화된 양의 단백질 복합체를 평형(equilibrium) 상태로 사용하여 수 개의 평행 측정(parallel determination)을 수행하도록 변형시킨, 실시예 3의 방법을 적용시켜,
- [0226] (b) 상기 평행 측정에서 얻은 정량적 정보를 기본으로, 다수의 단백질-파아지 상호반응에 대한 정량적 결합 곡선(quantitative binding curves)을 작도할 수 있으며 해리 상수와 결합능(binding capacity) 정보를 계산할 수 있다.
- [0227] 실시예 9
- [0228] 발명의 화학양론(Stoichiometry of invention)
- [0229] 결합 화학양론에 관해서, *E. coli*의 세포 용적(volume) 중 1 nM의 단백질은 대략 1개의 분자/세포이며 2,000개의 분자/포유동물(HeLa) 세포이고, 시그널화 단백질에 대해 특징적인 농도 (여기에서 예로서)는 10 nM 내지 1 마이크로M 범위이다. 또한, 파아지 디스플레이 항체의 해리 상수(Kd)가 10 nM에서 0.1 nM에 이르는 범위에 있고 분리속도(off-rate)가 $10^{(-3)}$ 내지 $10^{(-4)}$ s^{-1} 이며 이들 파아지가 일상적으로 선택될 수 있기 때문에, 대부분의 단백질/에피토프에 대해 포화 결합 화학양론이 예측되며 분리속도는 초기 해리없이 복합체가 구획화되기에 충분한 시간을 제공한다.
- [0230] 세균, 효모, 및 포유동물 세포 (Bioessays. 2013 Dec;35(12):1050-5.)에 입방 마이크론 (즉, 1 fL) 당 2 내지 4×10^6 개의 단백질이 있으며, 5000개의 진핵세포의 용적 (10000 fL)의 경우, 10^{10} 개의 단백질이 있고 최대 파아지 농도가 대략 $10^{16}/ml \rightarrow 10000fL$ 용적에: 10^{11} 개의 파아지이다.
- [0231] 상호반응체 복잡성은 10(+4-5) 범위이며, 목적하는 상호반응/파아지 중복화도(multiplexity)는 약 10으로, 이는 10⁵ 개의 개별 파아지, 즉 개별 파아지 당 0.1 nM (10^{16} 개/ml에서)에 상응하며, 10000 fL 당 10⁶ 개 이상의 모든 단백질이 1 nM보다 큰 농도를 가지기 때문에 평균 0.1-1 nM Kd의 항체 파아지 (HuCAL GOLD 나노몰보다 작은 확률: 30%) (J. Mol. Biol. (2008) 376, 1182-1200)는 50-5%의 포화도를 제공할 수 있다. 이는 25-0.25%의 동시-국소화된 포화에 상응한다 (동시-국소화된 포화란 다른 특이성을 갖는 2개의 결합된 파아지가 동일한 구획에 국소화되는 것을 의미한다).
- [0232] HiSEQ 2500 NGS 장치의 경우 0.25%의 동시-국소화 비율에서 3억 개의 판독치 (10-300 Gb, 250 bp 판독치)가 7×10^5 의 최소 개수의 헤테로다이머성 PCR 생성물에 상응하고, 이는 최소의 2원(binary) 상호반응 서열화 커버리지: 375임을 의미한다.
- [0233] 완전하게 랜덤화된, 1차, 높은 커버리지 항체 파아지 라이브러리의 복잡성은 $10^{(+13)}$ 의 각각의 파아지 클론에 달하며 관리가능한 에멀젼 PCR 또는 마이크로어레이 구획의 수는 $10^5 - 10^8$ 범위이고 (NGS 칩 구획의 현재 개수), 이러한 반비례수는 항체 파아지 라이브러리의 복잡성을 감소시켜 합하고 표적화된 (부분적 또는 전체 상호반응 체를 표적화) 단백질에 대한 결합능력을 갖고 있는 파아지의 존재비를 증가시킬 필요가 있다. 복잡성을 감소시키기 위하여, 특이적인 선택 공정(specific selection process)이 고안되었다 - 라이브러리 (cDNA)에 대한 라이브러리 (항체) 선택법을 사용하여 완전하게 랜덤화된 라이브러리로부터 파아지를 선택하는 방법으로 복잡성이 낮고, 친화성이 강화된, 천연의, 일반적인 목적의 파아지 라이브러리가 생성되고, 또한, 이 공정은 선택 공정 중에 항체-단백질 결합을 검출함으로써 모니터할 수 있거나 예를 들어 상이한 양의 투입 단백질 디스플레이 파아지를 사용하여 결합 운동역학 정보도 뽑아낼 수 있다.
- [0234] 또한, 상향식 방법(bottom-up) (결합 특성이 알려져 있는 파아지를 혼합하고 배경 파아지를 가함; 이들은 특정 목적으로 재단된 특이적인 라이브러리이다), 및 하향식 접근법 (예를 들면, 상호반응하는 파아지를 선택함으로써 복잡성을 감소시킴)을 사용하여 더 많은 양의 복합체 라이브러리를 점진적으로 구축함으로써 라이브러리를 생성시킬 수 있다.
- [0235] 실시예 10
- [0236] 수득한 정보의 통계적 평가(Statistical evaluation of the information gained)
- [0237] (a) 표준 방법으로 단백질 복합체를 추출하고, 항체 라이브러리 파아지를 합하여 형성된 복합체를 회석시키고 구획화시켜 비결합 파아지에 대해 단일 파아지 수준으로 분리시킨다.
- [0238] (b) 구획 당 파아지 (결합 및 비결합)의 항체 코딩, 특이적 DNA 단편을 증폭시키고 PCR (바람직하게는 제한된

수의 증폭 사이클)로 함께 연결한다.

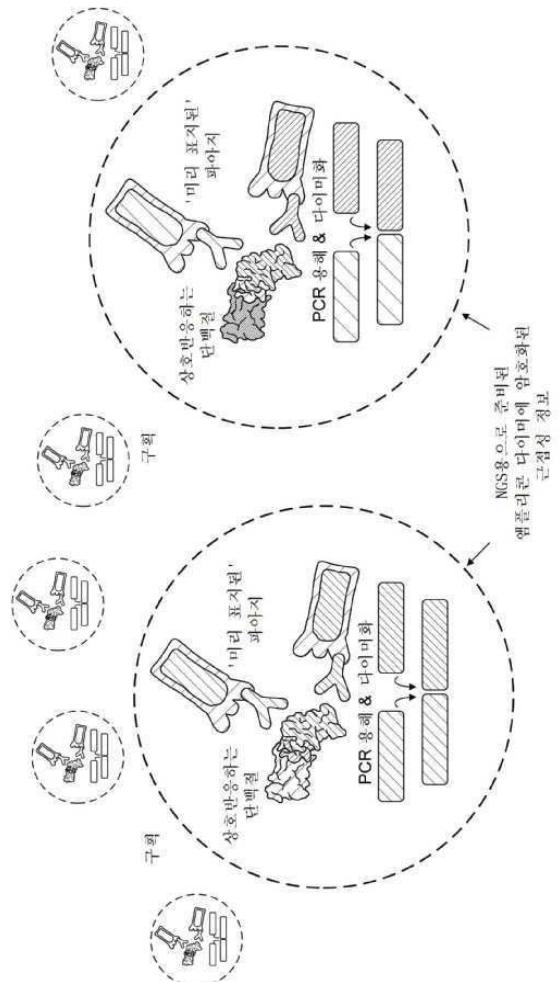
[0239] (c) 상기 연결 정보는 고도로 평행이고 정량적인 방식으로 차세대 DNA 서열화에 의해 밝혀진다.

[0240] (d) 각 파아지의 미리 측정된 결합 특성 정보와 결합된 파아지의 연결 정보를 사용하여 상호반응체를 컴퓨팅한다.

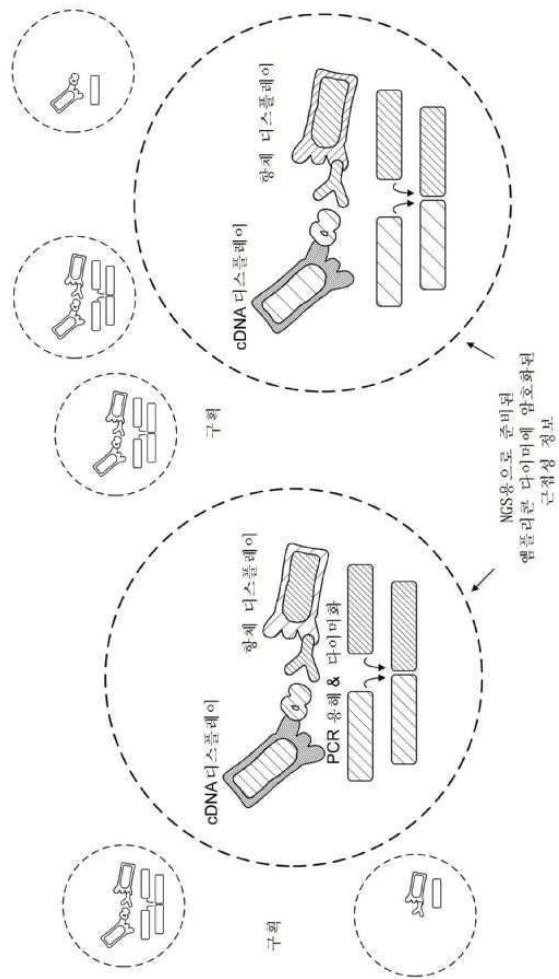
[0241] 본 방법은 단백질/단백질 복합체의 구획 기반 식별을 기본으로 하며, 여기서 구획 당 단일 단백질 복합체의 단백질-항체 동일성이 DNA로 번역된다. 연결된 모든 DNA 단편의 식별을 이용하여 상호반응을 정량적으로 측정하지만, 동일한 구획에서 돌발적으로 결린, 비결합 파아지는 배경에 기여할 수 있다. 이런 배경은 특이적인 경우와 구별될 수 있는, 랜덤한 경우에서와 같이 단순한 통계적 수단으로 조정될 수 있다. 구획화 중 결합체의 분포는 포아송 분포에 의해 지배되며, 따라서 구획의 수가 알려져 있는 경우, 각각의 결합체의 발생을 계수하고 (제한된 구획 기반 PCR 증폭 후 NGS에 의해 결합체의 상대적 존재비를 측정함으로써), 이에 따라 단백질/단백질 복합체의 배경 검출을 계산할 수 있다. 다수의 특이적인 결합 반응을 사용하여 다중-단백질 복합체 연결 정보가 동시-국소화에 기인하고, 직접적인 결합을 디스플레이하기 때문에 항체의 정확한 표적 단백질이 식별된다. 각각의 단백질/단백질 복합체의 경우 배경 검출이 계산되기 때문에 단백질/단백질 복합체의 검출방법의 임의의 변형은 실제적인 결합 효과에 기인하며, 이는 단백질/단백질 복합체의 계산된 배경 검출을 제거함으로써 단순한 뱃셈(또는 단백질/단백질 복합체 결합된 결합체가 결합체의 총수를 변화시키기 때문에 포아송 수정된 뱃셈)에 의해 계산될 수 있다. 평형 상태에서, 반응하는 항체와 단백질 분석물의 변화된 조합이 사용될 경우 Scatchard 플롯 또는 기타 결합 운동역학 계산식을 도식하여 항체 단백질 상호반응에 대한 Kd 또는 기타 파라미터를 계산할 수 있다. 상이한 농도의 상호반응하는 단백질을 사용하여 (실험 조건을 변화시키거나 스파이킹된 분석법을 사용하여) 모든 또는 수 개의 상호반응에 대한 상호반응체 내부 운동역학 데이터를 또한 계산할 수 있다.

도면

도면1



도면2



서열 목록

SEQUENCE LISTING

<110> GeneVillage Kft

Hart , Deborah

<120> Method

<130> P61471W0

<150> GB 1420852.4

<151> 2014-11-24

<160> 31

<170> PatentIn version 3.5

<210> 1

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> M13K07 Specific forward
<400> 1
ctcaagtccg tgacggtgat 20
<210> 2
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> M13K07 Specific Reverse
<400> 2
gacaaaaggg cgacattcaa 20
<210> 3
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> pBluescript II SK(+) specific forward
<400> 3
tcttgatccg gcaaacaac 20
<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> pBluescript II SK(+) specific reverse
<400> 4
ttttctgcgc gtaatctgct 20
<210> 5
<211> 18
<212> DNA
<213> Artificial Sequence
<220><223> pBluescript II SK(+) specific probe FAM-MGB labeled
<400> 5
ctggtagcgg tggtttt 18
<210> 6
<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> M13K07 specific probe VIC-MGB labeled

<400> 6

ccgtcaatat ttaccttccc 20

<210> 7

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> M13K07 specific forward Primer

<400> 7

taacgtggga atgggtcttc ctcaagtcgg tgacgggtat 40

<210> 8

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> M13K07 specific reverse primer

<400> 8

gacaaaaggg cgacattcaa 20

<210> 9

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> pBluescript II SK(+) specific forward

<400> 9

gaaggcaccat tcccacgtta tcttgatccg gcaaacaac 40

<210> 10

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> pBluescript II SK(+) specific reverse

<400> 10

ttttctgcgc gtaatctgct 20

<210> 11
<211> 18
<212> DNA
<213> Artificial Sequence
<220><223> pBluescript II SK(+) specific probe FAM-MGB labeled
<400> 11
ctggtagcgg tggtttt 18
<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> M13K07 specific probe VIC-MGB labeled
<400> 12
ccgtcaatat ttacccccc 20
<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> pBluescript II SK(+) specific reverse
<400> 13
ttttctgcgc gtaatctgct 20
<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> M13K07 specific reverse
<400> 14
gacaaaaggg cgacattcaa 20
<210> 15
<211> 18
<212> DNA
<213> Artificial Sequence
<220><223> pBluescript II SK(+) specific probe FAM-MGB labeled
<400> 15

ctggtagcgg tggtttt	18
<210> 16	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> M13K07 specific probe VIC-MGB labeled	
<400> 16	
ccgtcaatat ttaccttccc	20
<210> 17	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> General Forward Primer	
<400> 17	
ccaagaacac gctgtatctg ca	22
<210> 18	
<211> 40	
<212> DNA	
<213> Artificial Sequence	
<220><223> Dimerisation capable general reverse primer	
<400>	
> 18	
tgccatcca tttagaggt gagacggta ccagggtcc	40
<210> 19	
<211> 40	
<212> DNA	
<213> Artificial Sequence	
<220><223> Dimerisation capable general reverse primer	
<400> 19	
acctctacaa tggatgcgca gagacggta ccagggtcc	40
<210> 20	
<211> 21	
<212> DNA	
<213> Artificial Sequence	

<220><223> Dimer specific real-time forward primer
<400> 20
agttggagtc ttgggtcag g 21
<210> 21

<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Dimer specific real-time reverse primer
<400> 21
aggtgggtcg atgttgact actg 24
<210> 22
<211> 19
<212> DNA
<213> Artificial Sequence
<220><223> Dimer specific real-time PCR Probe
<400> 22
tctcacctct acaatggat 19
<210> 23
<211> 17
<212> DNA
<213> Artificial Sequence
<220><223> Anti-beta-galactosidase probe
<400>
> 23
gctagggcta tgtatcc 17
<210> 24
<211> 19
<212> DNA
<213> Artificial Sequence
<220><223> Anti-bovine ubiquitin PRobe
<400> 24
tgggtcgatg tttgactac 19
<210> 25
<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> General forward Source BioScience (6001_hDAb) library specific primer

<400> 25

ccaagaacac gctgtatctg ca 22

<210> 26

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Dimerisation capable general reverse primer Source BioScience (6001_hDAb) library specific primer

<400> 26

tgcgcattcca ttgttagaggt gagacggtaa ccagggttcc 40

<210> 27

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> General forward PhD12 Phage Display Peptide Library specific primer

<400> 27

cgcaattcct ttagtggtac cttt 24

<210> 28

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Dimerisation capable general reverse primer PhD12 Phage Display Peptide Library specific

<400> 28

acctctacaa tggatgcgca tctgtatggg atttgctaa acaact 46

<210> 29

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Dimer specific real-time forward primer

<400> 29

cggacgttg aaagttgtt agca 24

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Dimer specific real-time reverse primer

<400> 30

ggtcaccgtc tcacctctac 20

<210> 31

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> Dimer specific real-time probe

<400> 31

catacagatg cgcatcc 17