초파리 게놈의 3차원 접힘과 기능 발현 조직 원리

권 석 준

한국과학기술연구원 다원물질융합연구소

E-mail: sjoonkwon0531@gmail.com

요약문

염색체는 유전 정보가 물리적 구조로 환원된 형태이기 때문에 그로부터 정보를 읽고 전파하는 데 있어 기초가 되는 단위이다. 본 요약 페이퍼에서는 2012년 Cell지에 보고된 초파리 배아세포핵의 염색체에 대한 개량 3C 방법 (Hi-C)을 통한 대규모 염색실 시퀀싱 및 매칭 데이터 처리와 그로부터 구조적으로 재조합된 염색체의 하위 구획간 접촉 빈도에 관한 수학적 모형 연구와의 비교에서 발견한 사실을 요약 정리하였다. 원 논문에서는 게놈 단위의 개량된 염색체 형태 포착법(3C)을 통해 초파리의 배아세포핵의 고해상도 염색체 접촉 빈도 지도를 작성한 결과를 보고한다. 특기할 만한 사실은, 게놈 전반에 걸쳐, 물리적으로 경계가 뚜렷하게 나뉘는 구획이 후생 유전적 전사 기능을 촉진 혹은 저하시키는 부분과 매우 큰 상관관계를 이루며 분포한다는 것이다. 염색체 접촉은 그 내부의 물리적 구획간의 수직적 구조 형성 과정에 의해 결정되고, 접촉 빈도와 구획간 집속(clustering)을 설명하기 위한 전역 모형은 비활성화 구획이 염색체 내 특정 구역에만 한정되어 조밀하게 분포하는 반면, 활성화 구획은 염색체 전역에 뻗어 나가, 멀리 떨어진 구획간, 심지어는 서로 다른 염색체에 있는 구획간 장거리 접촉에 관여하고 있음도 드러났다. 이러한 원문의 결과를 종합해 보면, 초파리 염색체 접촉 빈도 지도에 기반을 두는 정량적 접촉 빈도 예측은 향후 인간 게놈의 염색체 구조 분석과 그 구조에 대응하는 유전 정보의 발현을 추적하는 연구에 있어 중요한 토대가 될 것이다.

Key Words: Drosophila, clustering, physical domain, Hi-C technique, contact map

> 본 자료는 Three-Dimensional Folding and Functional Organization Principles of the *Drosophila* Genome. *Cell.* 2012; 148(3): 458-472 의 논문을 한글로 번역, 요약한 자료입니다.

목차

1. 서론
2. 주요결과
   2.1 초파리 배아세포핵 내부의 염색체내 접촉 빈도를 표시하는 종합 지도

2.2 게놈은 물리적으로 뚜렷한 경계를 가진 구획으로 나뉠 수 있다
2.3 물리적 구획은 후생 유전적 구획을 반영한다
2.4 구획 경계는 절연체 접속점으로 결정된다
2.5 기능 저하적 물리적 구획은 뚜렷한 접힘 구조를 가지며 수직적으로 서로 상호작용 한다
2.6 구획 접촉 집속 분포도는 게놈에서 활성/비활성 부분이 차지하는 비율을 알려준다
2.7 기능적으로 특별한 장거리 접촉은 전역적인 접힘 원칙을 넘어 이루어진다
3. 토론
3.1 초파리 배아세포 핵의 게놈 전반에 걸친 염색체 접촉 지도
3.2 구획 기반의 염색체 수직 구조
3.3 물리적 구획은 결국 후생 유전적 구획에 대응한다
3.4 전역적 염색체 구조는 특수한 접촉을 위한 토대를 제공한다
4. 결론

1. 서론

실제 유전 정보를 물리적으로 전달하는 염색체는 세포의 핵 내부 특정한 위치에서 고도로 규칙성을 갖는 구조를 갖는다. 이러한 염색체 형태는 그에 대응되는 생명 현상 조절 기능과 맞물려 있을 가능성이 높은데, 이는 염색체가 후생 유전적인 조직 발현, 게놈 기능 조절, 다양한 상태의 세포 정보의 후생 유전 전달을 위한 일종의 매개체 역할을 하고 있을 가능성이 높음을 의미한다. 따라서 염색체 구조를 완벽히 파악하는 것이야말로 게놈 및 후생 유전 연구에 있어 필수적이고, 특히 염색체 접힘은 유전 정보의 암호화/해독과 어떤 인과관계가 있는지에 대한 해답을 제시할 것이다. 그간 염색체 섬유질 접힘을 설명하기 위해 다양한 수학적 모형이 제시되었지만, 보다 확실한 결론을 얻기 위해, 게놈 단위의 고해상도 DNA 조각간의 유전 정보 거리 측정, 상호 작용 빈도 측정이 절실히 요구되고 있다. 이를 위해 최근 활발히 쓰이는 3C (Chromosome Conformation Capture (염색체 형태 포착법))나 4C (Chromosome Conformation Capture on Chip (바이오칩 위에서의 염색체 형태 포착법)) 같은 기술은 염색질의 구조와 기능간 연관성의 세밀한 분석에 적극 활용되고 있다. 그럼에도 불구하고 아직 3C나 4C에서 얻어진 엄청난 규모의 데이터를 해석하여 염색질과 특정 게놈 기능 간에 어떠한 관계가 있는지에 관해 더 자세하게 이해하기는 요원한 상황 상황이다. 원 논문은 이러한 상황에 착안하여, 개량된 Hi-C 기법을 활용하여 초파리 배아세포핵의 염색체에 대한 고해상도 접촉 빈도 지도(Contact map)를 만드는 방법 그리고 그 의의에 대해 논하고 있다. 특히, 원 논문에서 발견된 가장 중요한 사실은 염색체가 기능에 따라 고도로 수직적 방식에 의해 구획 간의 상호작용이 가능해져 그 구조가 형성된다는 것이다. 이는 마치 단백질의 접힘이 여러 단계에 걸쳐 이뤄지고, 각각의 단계에서 필수적인 정보가 고차원적인 접힘 구조에 매칭되는 것을 연상케 하는 부분이다. 원 논문의 실험 결과 먼저 절연체 단백질이 특별하게 더 높은 농도로 분포하는 영역에서 구획간 경계가 결정되고, 다음에 이러한 구획이 뚜렷하게 접히면서 서로간에 수직적으로 상호작용한다. 구획을 크게 활성/비활성으로 단위를 나눴을 때, 비활성 구획끼리는 서로 뭉치고 특정 영역에만 집중 분포하는데 반해, 활성 구획은 덜 조밀하게 분포하면서

다른 활성화 구획과 상호작용 접촉을 한다. 이를 설명하기 원 논문은 간단한 확률 모형을 제시하고 이를 이용하여 염색체 접촉 빈도 지도 구조를 예측하여 활성화/비활성화 구획의 구조, 각각 혹은 서로간의 집속(clustering) 특성을 설명하려고 시도하였다. 그 결과, 모형은 전역/국부적인 접힘과 그에 연관된 접촉 빈도를 충분히 설명할 수 있는 것으로 나타났다. 따라서 원 논문이 제시하는 정량적 분석 데이터는 염색체 구조가 여러 단계에 걸쳐 게놈 기능과 밀접하게 연관된 것임을 보여준다고 할 수 있겠다.

2. 주요결과

2.1 초파리 배아세포핵 내부의 염색체내 접촉 빈도를 표시하는 종합 지도

원 논문의 저자들은 Hi-C 방법으로 초파리 배아세포핵 게놈 전체에 걸쳐 접촉 빈도를 정량 분석하였다. 이때 접촉 빈도는 염색체 내 염기쌍 간 거리와 대략 멱함수 관계(Power law)를 이루며 반비례하는 것으로 나타난다. 이전에 4C로 분석된 유전체 위치에 대한 접촉 지도와 비교해 보면 일단 Hi-C 데이터를 기초로 하는 접촉 빈도 분석 기법이 적절함을 확인할 수 있다

2.2 게놈은 물리적으로 뚜렷한 경계를 가진 구획으로 나뉠 수 있다

염색체 접촉 지도를 더욱 체계적으로 분석하기 위해, 간단한 접힘 원리와 염색체의 전역/국부적인 특성을 고려하여 가능한 한, 모든 구조를 가능성에 넣는 확률 모형이 원문에서 제시되었다. 첫 번째 단계로, 국부적 Hi-C 데이터에서 얻은 접촉 빈도에 다양한 노이즈를 넣는 모형이 제시되었고, 이 모형에서 두 영역의 접촉 확률은 영역간의 유전정보거리의 함수로 표현되었으며, 이 거리를 단위 재조정하면, 모형값과 실험값 간의 일치도를 확보할 수 있음을 보였다. 그 결과, 흥미롭게도 Hi-C 접촉 빈도 행렬의 대각성분을 따라 접촉 빈도가 높은 부분끼리 하위 행렬로 뚜렷하게 묶일 수 있음이 드러났다. 이는 실제 염색체내 구획의 물리적 경계와 연관된 것이다.

2.3 물리적 구획은 후생 유전적 구획을 반영한다

본 연구는 기존의 초파리 선형 표준 후생유전정보 종합지도와 접촉 지도에서 보이는 물리적 구획과의 상관 관계 분석까지 확장되었다. 그 결과, 403개의 선형 후생유전정보는 315개의 물리적 구획과 강한 상관성이 있는 것으로 나타났다. 실험 결과, 염색질 구획의 3차원 구조는 후생유전기능과 매우 강한 연관성이 있음을 알 수 있다.

2.4 구획 경계는 절연체 접속점으로 결정된다

염색질 분석 결과, 물리적 구획 경계 결정에는 여러 요소가 참여하는데, 이들은 모두 구획 경계에서 매우 조밀하게 분포함을 확인된다. 경계에서 무슨 일이 벌어지는지가 별로 중요하지

않다는 것은 아니지만, 경계 형성에 참여하는 접속점 비중은 상대적으로 낮은 편이다. 후생유전정보에 따라 수직적으로 집속되는 구획 경계에서는 이러한 절연체 조합이 반복적으로 이뤄지는 것으로 보이며, 이는 절연체가 접속되는 부분에서 구획 경계가 생긴다는 추론과도 부합한다. 특히 절연체들은 경계를 기점으로 거의 대칭분포 한다. 따라서 구획 경계는 절연체 단백질에 의해 결정됨을 알 수 있다.

2.5 기능 저하적 물리적 구획은 뚜렷한 접힘 구조를 가지며 수직적으로 서로 상호작용한다

원문이 제시하는 방법에 따라 유전 정보간 거리를 재설정하면 실험치와 상당히 잘 맞는 예측치를 보이는 모형을 만들 수 있다. 하지만 기능 저하적 후생유전 기능성 그룹에 속한 몇몇 구획의 경우, 유전 정보 거리와 접촉 강도의 멱함수 관계에서 다른 활성화 구획의 값에 비해 뚜렷이 다른 감소 지수를 보였는데, 이는 이들 구획의 접힘이 활성화 구획 접힘과 근본적으로 다르기 때문이다. 특히, 뚜렷한 접힘 경향은 수직적인 관계를 가지면서 보다 높은 수준에서의 질서도를 갖는 염색체 구조를 이룬다는 사실이 매우 중요하다. 기능 저하적 구획쌍에서 발견되는 일정한 접촉 빈도 분포는 이들 구획이 염색체 내부에서 수직적 구조 형성을 촉진시키고, 그 과정은 3차원 접힘과 관련되며, 또한 접촉 구획끼리의 장거리 작용에 의해 이뤄짐을 의미한다.

2.6 구획 접촉 집속 분포도는 게놈에서 활성/비활성 부분이 차지하는 비율을 알려준다

Hi-C 데이터에서 얻을 수 있는 구획끼리의 장거리 상호 작용에 대한 정보는 앞서 언급한 간단한 모형만으로 설명하기에는 역부족이다. 이를 위해 전체 활성/비활성 염색체 성분끼리의 물리적 집속 관계에 기반한 다른 모형을 제시해 본다. 이 모형에서 말하는 집속이란 사실 활성/비활성 구획에서 특별히 이에 관련된 부분끼리의 선택적 짝 형성에 기반한 것일 수도 있다. 원문에서는 이러한 추정이 맞는지 알아 보기 위해, 구획 단위의 접촉 빈도 행렬을 만들고, 이를 크게 두 그룹으로 구분했다. 그 결과, 활성화 구획끼리의 장거리 상호작용 빈도는 높은 반면, 비활성화 구획끼리 혹은 활성화-비활성화 구획간의 접촉 빈도는 상대적으로 낮게 나타났다. 따라서 비활성화 구획과 관련된 접촉은 특정 영역에서만 발생하고, 서로 다른 염색체 구역간의 접촉은 빈도가 낮아진다. 반대로 활성화 구획의 경우, 원래 속한 염색체 영역 밖으로 뻗어나가, 서로 다른 염색체 간 장거리 접촉에도 관여하여, 기능적으로 별다른 연관성이 없는 다른 활성화 구획과 접촉할 수 있게 되는 것으로 보인다.

2.7 기능적으로 특별한 장거리 접촉은 전역적인 접힘 원칙을 넘어 이루어진다

종합적으로 염색체간 접촉을 위한 수직적 구획 형성 모형에서, 유전 정보 거리, 물리적 구획경계 형성, 염색체를 전역에 걸쳐 활성/비활성 그룹으로 구분하는 것 등을 고려하면, 게놈 단위에서의 접촉 확률을 추정할 수 있다는 것이 원문이 제시하는 바이다.

3. 토론

3.1 초파리 배아세포 핵의 게놈 전반에 걸친 염색체 접촉 지도

원문에서는 게놈 단위의 개량 Hi-C 방법으로 염색체 접촉 지도를 제작하였다. 이를 통해 상당히 종합적이면서도 정확하게 초파리 배아세포핵 염색체 구조를 파악하고 접힘이 어떻게 이뤄지는지 체계적으로 보였다.

3.2 구획 기반의 염색체 수직 구조

원문에서 활용된 단순한 모형이 어떻게 10만배 이상의 접촉 빈도 차이를 보이는 대규모 데이터를 고려하며 실측치와 모형 예측치 간의 차이는 모형 복잡도와 어떤 관계를 갖는지 분석한 결과로 얻은 중요한 사실 중 하나는, 염색체가 물리적 구획으로 구분될 수 있다는 것이고, 이는 접촉 지도 행렬상에서 대각 방향으로 접촉 빈도가 최대값 분포를 갖게 구분된 하위 행렬로 나타난다 (그림 2E. 3R 염색체 부분에 대한 정규화값으로 표시된 접촉빈도 지도, 그림 3F. 3R 염색체 부분의 Hi-C 데이타에 근거한 정규화값으로 표시된 접촉빈도 지도). 모형에 따르면, 이러한 구분 형태는 단순히 접촉 자체의 조밀성이나 접촉에 관여하는 국부 구조만으로는 설명이 안 되며, 염색체내 접촉 확률 분포로 설명할 수 있다고 한다 (그림 5A-5C. 3R 염색체 부분에 대해 수정된 접촉빈도 지도 및 제시된 모형에 의해 이론적으로 예측되는 접촉빈도 지도와의 비교). 또한 물리적 구획이 염색체의 수직적 구조를 이루는 기초 단위가 되고, 이는 게놈 요소간 접촉 빈도가 물리적 구획의 위치가 아닌 성질에 의해서만 결정되는 데에서도 볼 수 있다. 그 결과, 염색체의 수직적 구조 형성이 가능해지고, 이러한 메커니즘은 후생 유전 조직의 기능과 연결되는 패러다임 입장에서는 상당히 긍정적인 정보일 수 있다

3.3 물리적 구획은 결국 후생 유전적 구획에 대응한다

물리적 구획은 염색체 접힘의 가장 기초를 이루는 기본 단위이고, 상위 수준에서는 후생 유전 정보의 발현과 인과관계를 가지고 있음이 거의 확실하다. 대부분의 경우, 절연체 단백질이 구획 경계를 뚜렷하게 만드는 역할을 한다.

3.4 전역적 염색체 구조는 특수한 접촉을 위한 토대를 제공한다

염색체 접힘은 매우 복잡한 방식으로 그리고 서로 다른 영역에 걸쳐 일어난다 (그림 2A. 3R 염색체 부분에 대한 광역 접촉 빈도 지도). 이 과정에서 장거리 염색질 간 상호작용이 어떤 의미를 가지고 있는지 알아 보려면 먼저 염색체 속의 섬유질 접힘을 정의할 수 있는 기준이 있는지, 있다면 그러한 기준을 지배하는 원칙이 있는지 이해하는 것이 필수적이다. 이러한 면에서 Hi-C 데이터는 염색질 접힘과 그를 설명하기 위한 모형 간의 가설 검증을 위한 중요한 정보를 제공한다 (그림 7A.

수직 구조 구획 모형). 우선적으로 원문 그림 2C (그림 2C. 유전적 거리를 보여주는 개략도)에 나타내었듯, 본 연구에서는 접촉 빈도에 이질적인 요소가 보였음 것을 확인하고 그 배경을 설명하였으며, 이를 통해 물리적 염색질 구획이 접힘에 있어 어떻게 배치되는지를 파악할 수 있었다. 둘째로, 이러한 물리적 구획은 염색체간 접촉 빈도에 따라 두 집속 단위 (활성화 그룹과 비활성화 그룹)로 나뉠 수 있고, 이 과정에서 두 집속 단위간, 혹운 각 집속 단위끼리의 접촉이 어떻게 이루어 지는지 설명할 수 있었다 (그림 6A. 3L 염색체 부분의 조립질 구획에 대한 정규화값으로 표시된 접촉빈도 지도).

4. 결론

요약하자면, 원문에서 보고된 Hi-C 실험 결과는 염색질 간 상호작용을 나타내는 종합 지도 구축에 필수적인 기초 자료를 제공하면서 염색체 구조와 실제 기능 발현 간의 상관관계를 파악하는 중요한 토대를 제공한 것으로 보인다. 물리적 구획의 접힘과 수직적 구조 형성 과정에 대한 분석은, 또한 기존에 연구가 많이 되지 않았던 염색체의 위상구조에 대한 정보도 제공해 준다. 원문에서 보고된 연구를 포함한 향후 이와 관련된 Hi-C 실험과 더 다듬어진 염색체 접힘 수학 모형은 핵 내부 구조와 형태 그리고 그에 의해 결정되는 게놈 기능 발현이 어떻게 연결되는지에 대한 더 많은 자료를 제공할 수 있을 것으로 전망한다.

The views and opinions expressed by its writers do not necessarily reflect those of the Biological Research Information Center.

권석준(2014). 초파리 게놈의 3 차원 접힘과 기능 발현 조직 원리. BRIC View 2014-R08
Available from http://bric.postech.ac.kr/myboard/read.php?Board=report&id=2218 (Jul 15, 2014)

Email: member@bric.postech.ac.kr