

## 제 10 회 137 억년 우주의 진화 4 강 노트- Epigenetics

(박문호 박사님의 강의를 초록한 내용입니다.)

전체 스코프를 잘 봐야 한다. 전체를 먼저보고 세부적인 것으로 들어간다.  
이 그림 하나만 철저히 이해하면 된다.  
알아야 할 것은 딱 2 가지뿐이다.

nucleotide 와 peptide 2 가지이다.  
nucleotide and peptide that's all!  
nucleotide 가 DNA 이고 RNA 이다.

peptide 는 아미노산이 2-3 개 연결된 것이고 이것이 확장된 것이 단백질이다.

세포 속에는 2 가지 밖에 없다. nucleotide 와 peptide.  
둘 다 tide 로 산성이다.  
DNA, RNA 묻지 말고 nucleotide 를 알아야 한다.  
nucleotide 는 입자물리학에서 보손과 페르미온 같은 개념이다.  
양성자, 중성자, 전자는 페르미온이고 w 입자, 포톤, 감마는 보손인 것과 같다.

집합관계 용어에 익숙해져야 헛갈리지 않는다.  
세포 속에는 딱 2 가지 밖에 없다.  
peptide 와 nucleotide 두 가지 용어를 정확히 알면 모두 해결된다.

손자 뺨을 잡았더니 할아버지 이었다. 공부가 묘해지고 헛갈린다. 손자가 할아버지다.  
모두 허상이고 실체는 한 가지다. 하나 밖에 없다. RNA 뿐이다.

NT(nucleotide)속에는 유전에 관한 모든 것이 있다.  
40 억년 전에 출현하였으며 종류는 수십 종이다.  
오늘 하나만 가져간다. 알아야 할 실체는 NT, 그 속에 유전, 진화 모든 것이 들어있다.

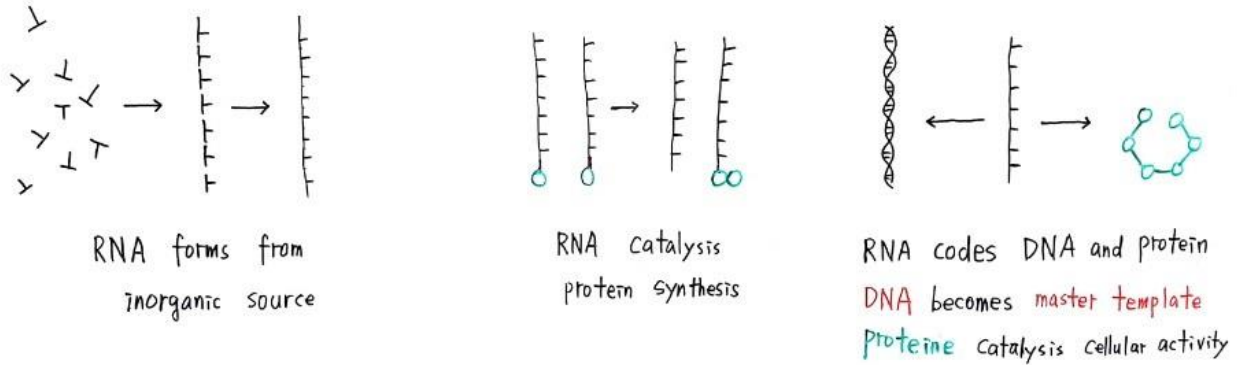
nucleo 는 핵이고, tide 는 산이다.  
핵 속에 있는 산성물질이란 뜻으로 1900 년대 초에 발견되었다.

그 후 1950 년대 초에 DNA 구조를 알게 되었다.  
RNA 에는 10 여가지가 있다. 초기에는 RNA 한가지 한가지를 잘 몰랐다. 모르니까 생물학에 구멍이 생겼다.  
생물학 혁명은 RNA 에서 일어나고 있다. 지금 생물학의 주축은 RNA 에 가 있다.  
그 모든 것을 담고 있는 것이 NT 이다.

분자를 보면 이빨 밖에 없다. RNA 10 가지 모두 이빨(NT)로 되어 있다.  
NT 가 30 억개 연결된 것이 DNA 이다. 수십 개에서 수천 개 모인 것이 RNA 이다.  
모두 NT 로 구성되어 있다. DNA, RNA 모두 NT 로 되어 있다.

세포 속에는 두 가지 물질이 있다.  
하나는 NT 이며 NT 가 모인 것이 RNA 와 DNA 이다.  
두 번째는 아미노산으로 아미노산이 모인 것이 단백질이다. 따라서 세포의 실체는 NT 와 아미노산이다.  
아미노산이 10 개 미만이면 peptide 이고, 수십에서 수천 개가 모이면 단백질이다.  
아미노산과 DNA 개념도 잊어라  
NT 와 peptide 만 알면 된다.

## RNA World



RNA 가 40 억년 전에 먼저 출현했다.

### RNA forms from inorganic source.

처음에 NT 가 있었다.

서로 인터랙션 하다 보니 일직선으로 배열되었다.

그 사이를 인산이 결합 시켰다. 인산이 5 프라임과 3 프라임을 결합시킨다. diester bond 가 생겼다.

여기서 NT 는 DNA 일수도 있고 RNA 일 수도 있으나 지금은 DNA 는 아니고 RNA 이다.r

모든 것은 NT 가 모여서 된 것이다.

RNA 는 coding 과 non coding RNA 가 있다.

non coding RNA 중에는

20 여개 모인 것이 small non coding RNA 이고

200 개 보다 작은 것이 small non coding RNA,

200 개 보다 큰 것이 long non coding RNA 이다.

mRNA 는 coding RNA 이다.

### RNA catalysis protein synthesis

10 억년 후 엮인 RNA 끝에 구슬(아미노산)이 붙었다.

하나의 RNA 에 여러 개의 구슬이 붙었다.

구슬이 모여서 단백질이 출현했다.

polypeptide 가 되었다. RNA 가 단백질 합성을 촉매 한 것이다.

### RNA codes DNA and protein

원래 RNA 는 single strand 이다.

RNA 외 가닥이 다른 RNA 외가닥을 만나 2 중 strand 변이가 일어나면 DNA 가 된다.

DNA 는 RNA 의 자손이다. 부모는 RNA 이다.

DNA 와 protein 이 RNA 에 의해서 만들어 졌다.

### DNA becomes master template.

DNA 는 그렇게 중요하지 않다. 그 동안의 DNA 개념을 씻어야 아버지 RNA 가 보인다.

DNA 는 손끝 하나 까딱하지 않는다. DNA 는 바뀌면 안 되는 운명을 타고 났다.

DNA 의 본질은 가만히 있는 것이다. 35 억년 동안 가만히 있었다.

손 까딱하면 사고 친다. 사고 치면 끊긴다. DNA 의 염기는 평균 1 억개이다. 끊어지면 안 된다. 수선에 엄청난 에너지가 든다.

DNA 본질은 먼지 뒤집어 쓰고 가만히 있는 것이다.

DNA 는 손끝도 까딱하지 않고 수동적으로 당한다.

RNA 는 수천 개가 넘어가면 분해된다.

DNA 는 가만히 있도록 만들어 졌다. 움직이면 안 된다. 수동적 돌연변이만 인정되었다.

DNA 는 팔만대장경 목판 같다. Template 이다. 글자가 바뀌면 안 된다. DNA 가 바뀌면 생명은 Chaos 가 된다. 수동적 돌연변이만 허용된다. 그 돌연변이를 통해서 수 백만종의 생명체가 생겼다.

Protein catalysis cellular activity

결국 RNA 가 다 했다. 이것이 RNA World 이다.

알아야 할 것은 NT 와 peptide 이다. 이 두 가지만 알면 된다.

RNA forms from inorganic source.

RNA catalysis protein synthesis

RNA codes DNA and protein

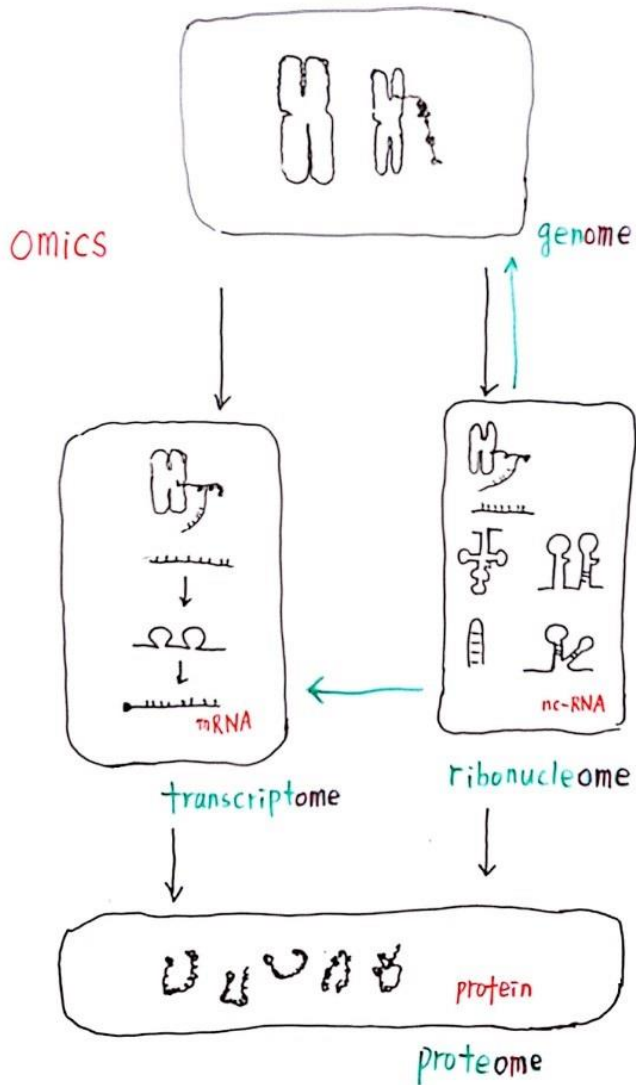
DNA becomes master template.

Protein catalysis cellular activity

헌법전문에 해당한다.

모든 생물학은 이것을 풀어 쓴 이야기다.

염색체는 세포 분열 시 DNA 이중나선이 실타래처럼 모인다.



Transcriptome is the set of all [RNA](#) molecules, including mRNA, rRNA, tRNA, and other non-coding RNA, produced in one or a population of cells.

Proteome is the entire complement of [proteins](#), including the modifications made to a particular set of proteins, produced by an organism or system.

유전 정보를 다 모은 것을 genome 이라 한다.  
 즉 인간 23 개 염색체 전부를 합친 개념을 genome 이라 한다.  
 이 속에 탄생에서 죽음까지 모든 것이 있다.

전사를 해서 나오는 전체 RNA 집합을 Transcriptome 이라 한다.  
 10 만개 오더이다.

non coding RNA 가 주역이 되고 있다.  
 하나의 염색체에는 NT 가 약 1 억개이다. 인간 23 개를 모두 연결하면 1.8m 나 된다.  
 염색사를 풀면 다 풀린다. 즉 양 끝이 있다. 인간 genome 은 직선이다. 박테리아는 원형으로 되어 있다.  
 인간이 죽는 이유도 게놈이 선으로 되어 있고 끝이 있기 때문이다. 세포 분열하면서 telomere 가 짧아져 죽게 된다.  
 박테리아는 원형으로 되어 있기 때문에 영원히 산다.

단백질이 코딩 되는 부위는 1.5% 밖에 되지 않는다. 98%는 non coding 이다.  
 지금은 non coding DNA 분야가 주역으로 떠오르고 있다.

어쩌면 non coding DNA 의 복권은 당연하다. 옛날의 시각은 틀렸다.  
tRNA, rRNA, miRNA siRNA 등이 모두 non coding DNA 영역에 있다.  
암과 관련된 유전자도 junk 영역에 있다.

실제로 junk 영역은 1.5%가 단백질을 코딩 하는데 도움을 준다. 효율을 높인다.

transcriptome 은 DNA 가 RNA 로 바뀌는 것과 관련된 모든 것이다.  
한 세포 또는 세포 집단에 있는 모든 RNA 분자의 집합이다.

그림에서 공통된 부분은 -ome 이다.  
OMICS 라 한다.

그림의 세 개 계층 모두를 다루는 학문을 OMICS 라 한다.  
결국 생물학은 OMICS 이다.

다시 말하지만 알아야 할 것은 nucleotide 와 peptide 이다.  
poly peptide 는 단백질이고  
poly nucleotide 는 핵산이다.

두 그림으로 생물학 다 끝난다.

ribonucleosome 이 genome 과 transcriptome 에 영향을 미친다.  
ribonucleosome 이 역으로 genome 에 영향을 미친다.

### 크리스퍼 가위

CRISPR 은 RNA 와 CAS 9 이 결합된 것이다.  
CRISPR 은 DNA 를 자른다. 유전병 치료에 혁명적이다.

알파고 만큼 중요한 것이다.

1 세대 가위는 zinc finger nuclease 이고 2 세대 가위는 TALENS(transcription activator like effector nuclease) 였다..

1 세대 가위는 4 개의 NT 를 매치 시켜야 했다. 즉 4 개의 순서가 동일한 NT 를 1 억개 염기에서 찾아 매치 시키는데 동일한 염기가 나올 확률이 매우 많아서  
부작용의 위험이 너무 컸다.

2 세대 가위는 염기 수가 10 개 이었다. 확률은 줄어 들었지만 그래도 부작용이 문제 이었다.

크리스퍼는 염기가 20 개이다. 1 억개중 20 개의 동일한 순서가 될 경우는 거의 없다.

디자인하여 만들 수도 있고 자를 수 있다. 부작용도 거의 없다.

또한 1 세대는 1 년이 걸렸으나 크리스퍼는 1 시간이면 할 수 있다.

부작용은 줄고 시간도 절약할 수 있게 되었다.

크리스퍼 혁명이라 할 수 있다. 수백 가지 암 치료에 적용이 가능하다.

새로운 개념을 이해해야 한다. 패러다임이 바뀌고 있다.

### RIBONUCLEOTOME

크리스퍼 가위가 nc RNA 의 대표주자이다.

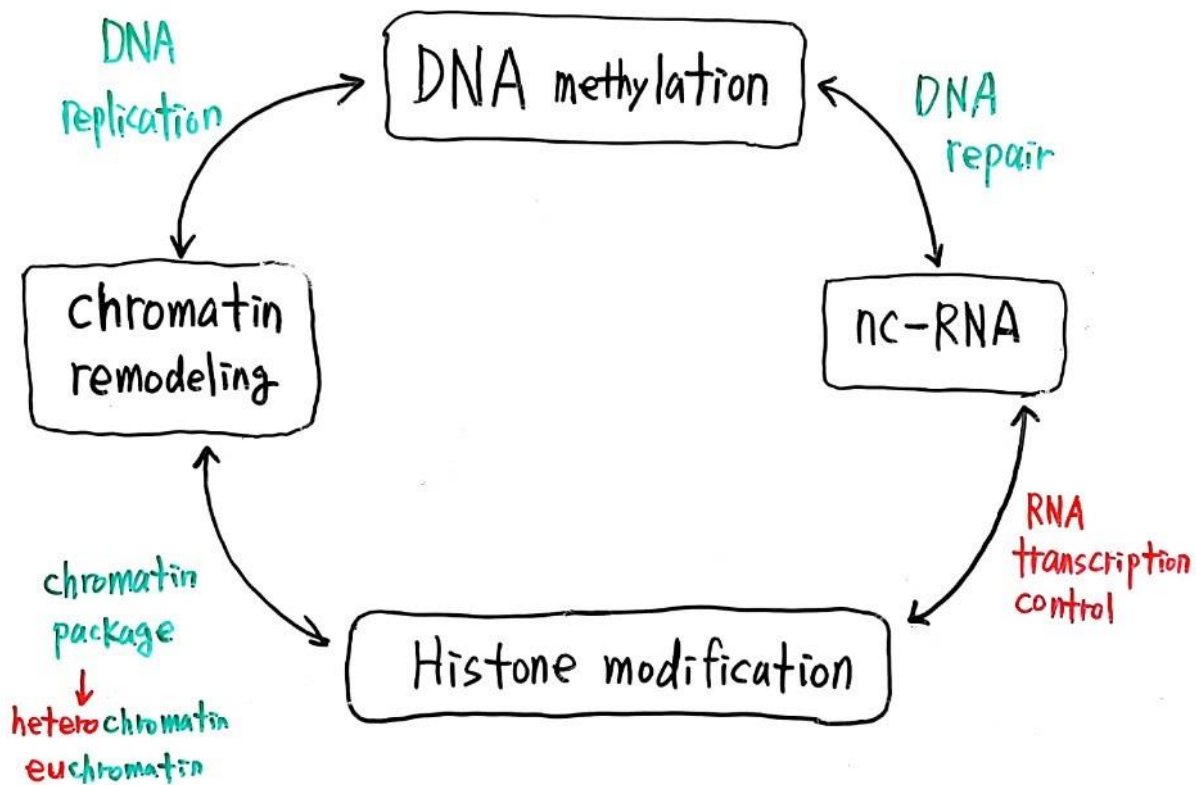
생명을 다시 보게 하는 놀라운 사건이다.

Epigenetics 는 이제 유전학을 초월하고 있다.

현재 까지는 mRNA 위주였다.

후성유전학은 출발부터 거의 대학원 과정이다 그만큼 어렵다.

후성유전학에는 4 개의 중심 개념이 있다.



## DNA methylation

### Histone modification

세포 안에 뉴클레오솜이 3 천만개나 있다. NT 가 뉴클레오솜을 1.65 바퀴 감고 있다. 뉴클레오솜을 감고 있는 NT 가 147 개이다. 뉴클레오솜과 뉴클레오솜 사이에는 NT 가 200 쌍 있다. 둘을 합쳐도 400 개가 되지 않는다. 그 사이에 폴리머라제와 여러 종류의 TF 가 붙는다. 메틸테이터도 있고 액티베이터도 있다. 너무 복잡하다. 박테리아는 1 초에 1000 개의 NT 를 심어야 한다. 히스톤이 3-4 개는 빠져야 한다. 빨리 풀어야 한다. 그래서 히스톤이 빠져 나온다. 또다시 감아야 한다. 대단한 동적 다이내믹스이다.

히스톤과 NT 는 무슨 힘으로 붙어 있는가.

히스톤은 단백질이다. 히스톤 단백질의 아미노산과 NT 사이가 수소 결합으로 붙어 있다.

NT 가 히스톤을 1.65 바퀴 감고 있는 동안 약 140 곳의 수소 결합이 있다. 수소 결합은 찌꺼기 같다. 70 도가 넘으면 떨어진다.

### Chromatin remodeling

진핵세포는 DNA 를 풀지 않는 방향으로 진화해 왔다.

단백질이 충돌할 수 없도록 된 구조를 heterochromatin 이라 하며,

단백질이 붙을 수 있도록 풀린 상태를 euchromatin 이라고 한다.

생명은 NT 와 polypeptide 와의 충돌 현상이다.

자연은 디지털이다. 아날로그는 없다. 인간의 메저먼트의 한계일 뿐이다.

폴리머라이즈는 모노머의 연결이다. 모노머가 디지털이다. NT 도 디지털이다. 그래서 끔거나 뺄 수 있다. 아날로그는 휘어진다.

NT 와 NT 사이는 자를 수 있다. 그 자르는 역할을 OH 기가 한다.

크리스퍼 가위가 DNA 를 자른다.

## Chromatin remodeling

CpG island 가 세포 하나에 100 만개 정도 있다. 프로모터 주위에 CpG island 가 있을 확률이 높다.

CpG 섬의 MAP 을 그리는 것이 핫 이슈이다.

시토신에 메틸기가 붙어 있으면 POL2 가 착륙이 불가하다. 메틸기를 탈락시키는 단백질이 작동 하고 난 후 착륙할 수 있다.

크로마틴 리모델링과 디메틸레이션이 동시에 일어나야 복제가 가능하다.

## DNA repair

DNA 메틸레이션과 nc RNA 가 함께하는 것이 DNA repair 이다.

DNA 는 고장이 많다. 주로 DNA 가 잘라진다. 잘라 지면 세린의 인산기가 즉시 잘린 부위를 붙인다.

크로마틴 리모델링과 히스톤 모디피케이션이 같이 하는 것이 크로마틴 패키지이다.

TF 와 중합효소가 앞으로 가기 위해서는 바위를 치워야한다.

헤테로크로마틴이 유크로마틴으로 바뀌어야 한다.

## RNA transcription control

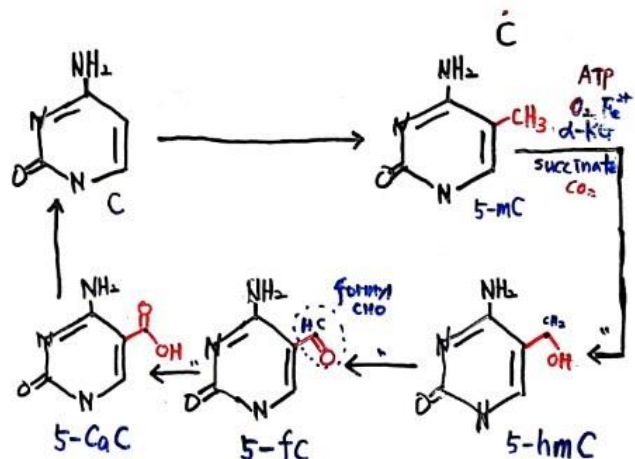
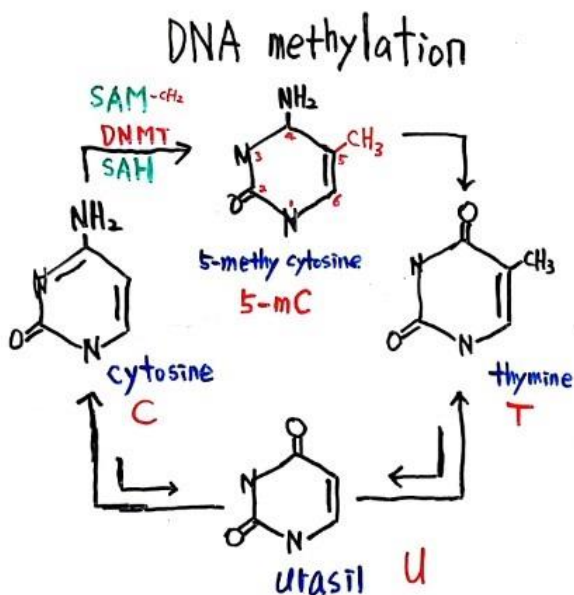
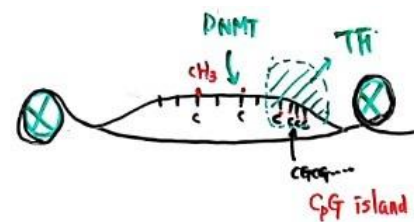
생명에서 가장 중요한 부분이다. 발생을 조절한다.

200 여 종의 세포가 생기고 식물과 동물 발생의 모든 이야기가 이 두 가지 사이에서 진행된다.

## DNA methylation

많은 내용이 있지만 다음 2 개의 그림만 알면 된다.

DNA 메틸레이션은 시토신에 메틸기가 붙는 현상이다.



공통되는 부분을 먼저 그린다. 피리미딘 구조는 60%가 동일하다.

피리미딘 분자식은 우라실 터 시작해야 한다.

피리미딘 생성시 Orotate 가 먼저 나오고 Orotate 에서 CO2 를 떼고 나온 것이 우라실이다.

우라실에서 시토신 그리고 티민이 나온다.

CpG island 의 시토신에 메틸기가 붙어 있는데 이 메틸기를 붙여 주는 단백질을 DNMT(DNA methyltransferase)라고 한다.

이 메틸기를 떼 주어야 TF 가 붙을 수 있다.

(동영상 참조)

SAM:S-adenosyl methionine

SAH:S-adenosyl Homosystein

DNMT:DNA Methyl Transferase

5-mc:5 methyl cytosine

5-hmc:5 hydroxy methyl cytosine

5-fc:5 formyl cytosine

5-CaC :5 carboxyl cytosine

$\alpha$  KG:  $\alpha$  ketoglutarate

후성 유전학을 왜 공부해야 하는가?.

우리가 먹는 음식, 읽는 책, 만나는 사람, 경험이 우리를 결정하고, 자손에게까지 영향을 미치기 때문이다.

유전-환경 논쟁은 오래 되었다. 2 분법으로 갈라져 50 년간 논쟁 했다.

그런데 이제 그 논쟁이 낡은 센스가 되고 있다.

환경이 유전자 자체는 바꾸지 못하지만 유전자가 발현될 확률을 바꾸고 있다.

Genetics 에 Dogma 가 있다. 그것을 건드리면 안 된다. 그것은 그대로 존재한다.

DNA 에 자외선이나 약물, 환경 변화에 의한 돌연변이가 있다. DNA 가 바뀌는 것은 돌연변이에 의해서만 바뀐다.

수정에 의해서만 유전자가 자손에게 전달 된다. 그 원칙은 바뀌지 않는다. 그것이 유전학이다.

그것은 100% 맞고 지금도 유지되고 있다.

뉴턴 역학이 일반 상대성 이론이나 양자 역학에 의해서 바뀐 것과 같다.

20 세기 들어 원자의 세계 등 미시세계에서 설명할 수 없는 현상들이 나타나기 시작하여

이제는 뉴턴 역학이 전체 역학의 부분 집합으로 남아 있다.

뉴턴 역학이 전체 역학에서 어떤 위치를 점하는가를 알게 되었다.

일반 상대성 이론과 양자 역학 속에서 뉴턴 역학이 재 정립되었다.

현재의 후성유전학은 30 년 전 유전학의 위치에 있다고 한다.

30 년전 유전학은 제네틱스를 밝히기 위해 많은 부분 부분을 보고 있었다.

지금 후성유전학도 그 위치에 있다.

모두 밝혀지면 생물학 교과서가 바뀌어야 한다.

뉴턴 역학이 양자역학에 의해 바뀌어 지듯이 유전학도 후성유전학에 의해 바뀌어 질 것이다.

지금 후성유전학은 시대를 앞서 간다. 삶을 바꿀 것이다.

암, 노화, 치매 등이 대부분 정크 유전자와 링크되어 있다.

그 동안 유전학은 이 부분에서 돌파구를 열지 못했다.

판을 바꾸는 게임이다. 새로운 영역이 생기는 게임이다.

100 년 전만해도 운동이 건강에 도움이 되는지 몰랐다. 지금은 70 억 인류가 너무나 당연한 사실로 받아들인다.

그 간단한 것도 전 인류가 이해한 것이 100 년도 안 되었다.

후성 유전학을 이해하면 삶을 바꾸고 모든 분야의 프런티어가 될 수 있다.

우리의 환경이 우리의 삶에 미치는 영향이 분자적으로 이해되고, 자손에게 전달되는 메카니즘이 설명된다.

메틸이 붙고 안 붙고는 환경에 달려 있다.물론 환경이 100% 자손에게 전달 되지는 않는다.

인간 유전자의 2% 정도에 메틸기가 붙어 있다.

정자와 난자의 DNA 에도 메틸기가 붙어 있다.

메틸기는 수정 시에 거의 리셋 된다. 그러나 그 일부가 전달된다.

유전자는 바뀌지 않지만 메틸기가 붙는 양상의 일부가 수정란을 통해서 자식에게 전달된다는 사실이 발견되었다.

그러면 우리가 좋은 환경에 살고, 좋은 친구 만나고, 좋은 책을 읽는 성향이 자식에게 전달 되면 ,



자식은 그 영역의 유전자가 발휘될 확률이 높아지고 성공의 가능성도 높아진다.

DNA 메틸레이션의 중요 개념은 다음 4 가지 이다.

DNA methylation → cellular pluripoteny  
chromatin compaction  
differentiation  
transcription control

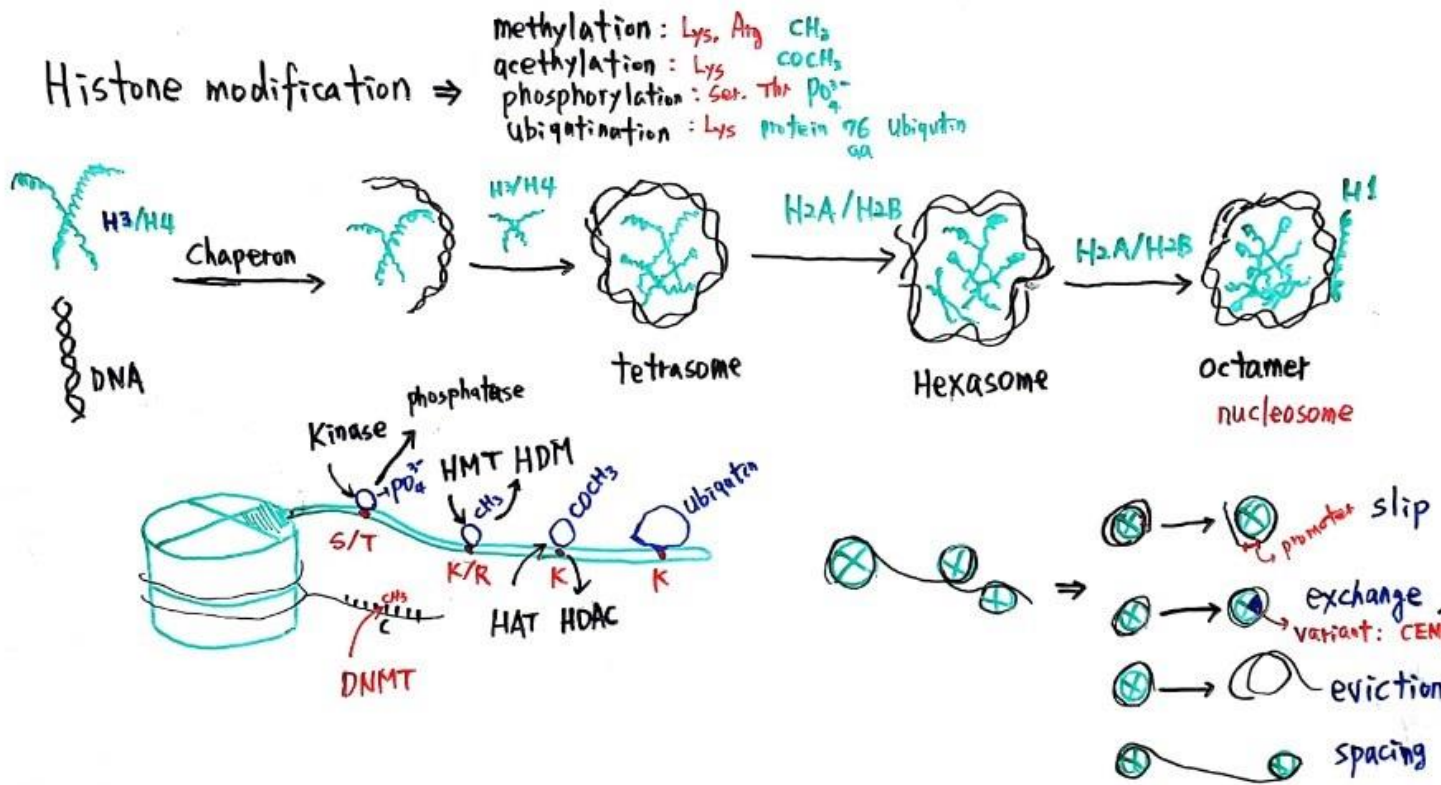
줄기세포 등의 다능성에 영향을 미친다.

DNA 메틸레이션에 의해 크로마틴의 형태가 영향을 받고 유전자 발현 여부가 결정된다.

분화 과정에 직접 관여한다.

CpG island 메틸화에 따라 전사조절인자의 활동에 영향을 미쳐 transcription 과정을 조절한다.

Histone modification



Chaperon 단백질의 역할로 히스톤을 DNA가 감는다.  
 팔량체가 될 때까지 1.65 바퀴를 감는다. 그러면 뉴클레오솜이 된다.  
 히스톤 단백질에 아미노산이 약 100 개 정도 연결되어 있다.  
 히스톤 단백질의 꼬리가 나와 있다. 꼬리도 아미노산이다.  
 히스톤 모디피케이션에는 4 가지가 있다.(동영상 참조)

methylation: 리신, 아르기닌- $\text{CH}_3$

뉴클레오타이드에 메틸기가 붙는 것은 DNA 메틸레이션이고, 아미노산에 메틸기가 붙는 것은 히스톤 메틸레이션이다.

acetylation: 리신- $\text{COCH}_3$

phosphorylation: 세린, 트레오닌- $\text{PO}_4^{3-}$

ubiquitination: 리신-protein 76 ubiquitin

인산을 붙이는 효소가 kinase 이고, 인산을 떼는 효소가 phosphatase 이다.

메틸기를 붙여주는 효소가 histone methyl transferase, 떼내는 효소는 histone demethyl transferase 이다.

마찬가지로 아세틸 기를 붙이는 효소는 histone acetyl transferase, 반대는 histone deacetyl transferase 이다.

그리고 시토신에 메틸기를 붙여 주는 것은 DNA methyl transferase 이다.

그리고 히스톤 모디피케이션에는 물리적 변화도 나타나는데 다음의 4 가지가 있다

slip: 프로모터가 있는 영역이 잘 나타나도록 미끄러진다.

exchange: 발생 등의 경우에 H3 히스톤이 CENP-A 로 교체된다.,

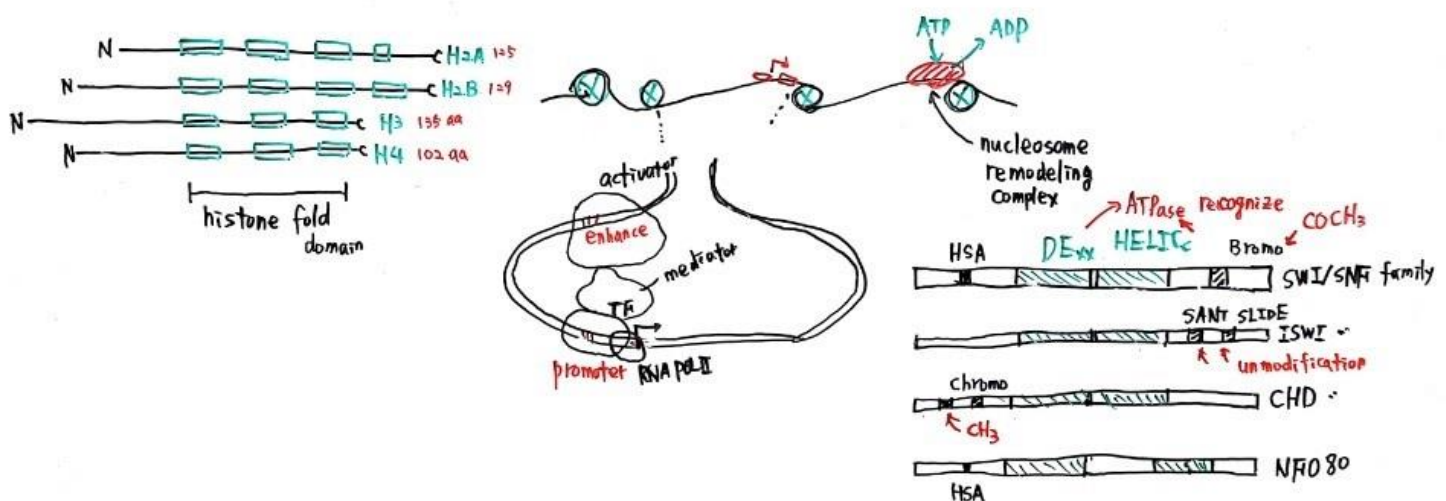
eviction: 히스톤이 빠져 준다.,

spacing: 뉴클레오솜의 간격을 조정

**H3K4**: euchromatin, TF↑  
**H3K9**: heterochromatin,  
**H3K27**: hetero- fault  
**H3K36**: eu- → Transcription elongation  
**H3K79**: eu- → main<sup>tain</sup> eu-  
**H4K20**: hetero, histone exchange

H3K4: H3 히스톤 4 번째 리신(K)  
 숫자를 기억하자 다음 시간에 더 깊이 들어간다.  
 후성유전학을 뿌리 뽑자.

Chromatin remodeling(동영상 참조)



기본 원리만 잘 알면 디테일을 풀어나가는 것은 어렵지 않다.  
 물론 세부적인 것을 암기해야 한다.

단백질의 종류는 100 만 종류가 되고, 하나의 세포 속에는 10 만개의 protein 이 있다.  
 protein 은 100 년 동안 연구되어 왔다.

뉴 클레오솜 리모델링은 ATP dependent nucleosome remodeling complex 단백질이 한다.  
 이 단백질을 분석해 보았다.  
 그랬더니 패미리를 구성하고 있었다. 4 가지 패미리가 있다. 각 각의 패미리는 서브 유니트로 구성된다.  
 아미노산 서열의 기능에 따라 분석할 수 있다.

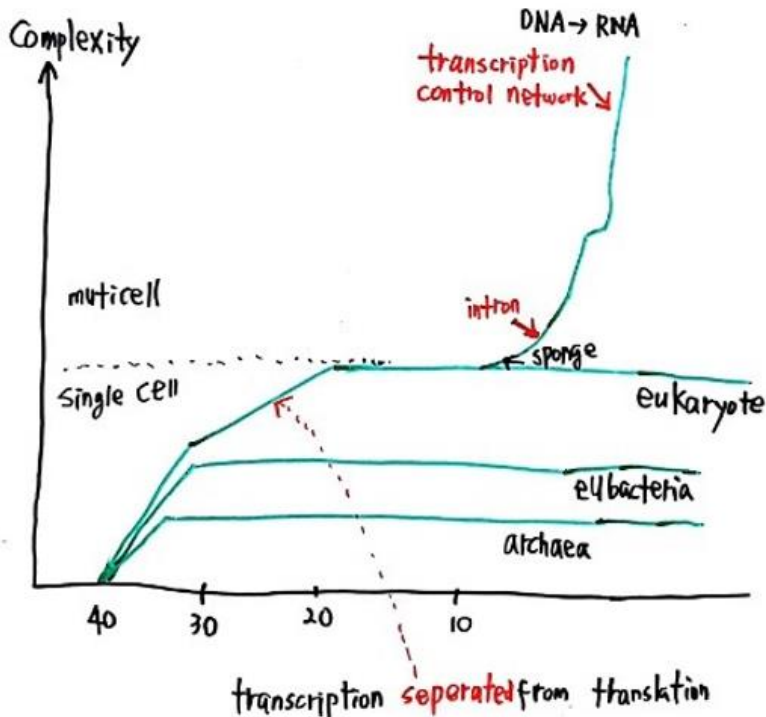
다음은 히스톤 단백질이다.

생물 중 마다 아미노산 갯 수가 다르다. N 말단과 C 말단을 구분해야 한다.

2002 년 네이처지 리뷰 페이퍼에 있는 도표이다.

nc RNA 가 왜 중요한지를 알게해 주는 그림이다.

후성유전학을 공부하는데 반드시 필요한 핵심도표이다.



지구 상 모든 생명체를 조사 했다.

생명의 진화를 보여 주는 도표이다.

가로 축은 년대를 표시한다. 세로 축은 Complexity(종의 복잡화)이다.

40 억년 전에 고세균이 출현하고 이어서 진정세균, 진핵생물이 출현했다

10 억년 전 쯤에 생명체의 복잡성이 증가하기 시작했다.

다세포 생물에서 제일 빨리 출현한 것이 해면동물이다.

30 억년 전 쯤의 복잡도 증가의 원인은 transcription separation from translation 이었다.

핵막의 출현과 관련이 있다.

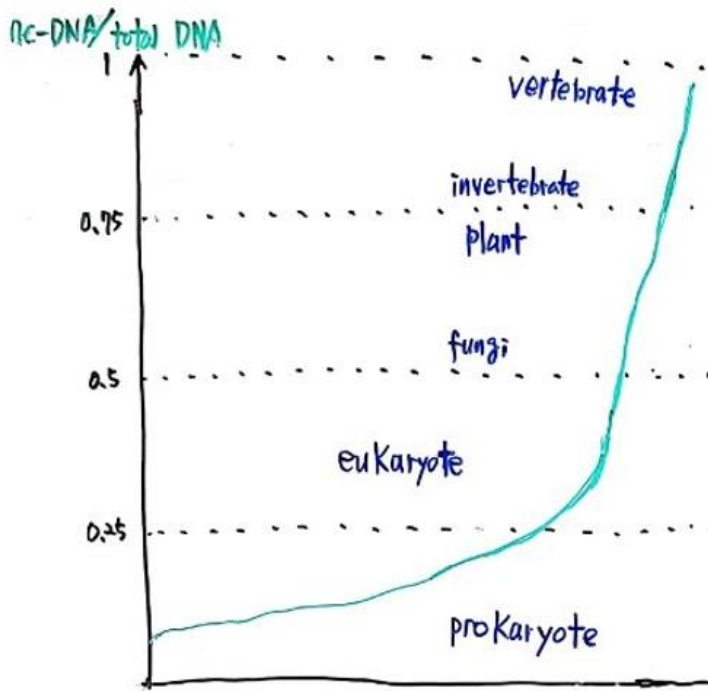
두번째 증가는 intron 이 개입되면서 생명 현상이 다양하게 되었다.

마지막 증가의 원인이 transcription control network(전사조절)의 출현이다.

이 혁명을 다른 말로 하면 epigenetics 와 같은 말이다.

DNA 에서 RNA 로 가는 과정이 엄청 다양해지고 RNA 세계가 꽃이 피었다.

전사 조절이 가장 핵심이었다.



생명체의 다양화와 유전물질과의 관계를 나타낸 도표이다.

세로축은 게놈 전체 DNA 에서 nc DNA 의 비율이다.

nc DNA 의 비율이 높을수록 더 복잡한 생명체가 되었다.

nc DNA 가 고등한 생명체를 만드는 장본인인 것이다.

다른 것과는 상관 관계가 약했고 nc DNA 와 상관관계가 높았다.

인간의 nc DNA 의 비율은 98%가 넘는다.

우리가 유전자를 보는 시각을 바꿔 놓았다.

유전자 갯수가 중요하지 않고 코딩하지 않는 부분이 얼마나 많은가가 더 중요하다는 것을 알았다.

junk DNA 를 연구하는 분야가 후성유전학이다.

그기서 nc RNA 가 가장 핵심이다.

ncRNA 속에 tRNA, rRNA, miRNA 모두 들어 있기 때문에 기존의 유전학도 다시 다 공부해야 한다.

그 속에서 재정립된다.

우리가 쓰레기라고 생각했던 것이 결국 보배 덩어리로 밝혀 졌다.

전모가 밝혀 지면서 우리가 그 동안 몰랐던 문제들에 어떤 식으로 접근해야 될 것이지 밝혀지고 있다.