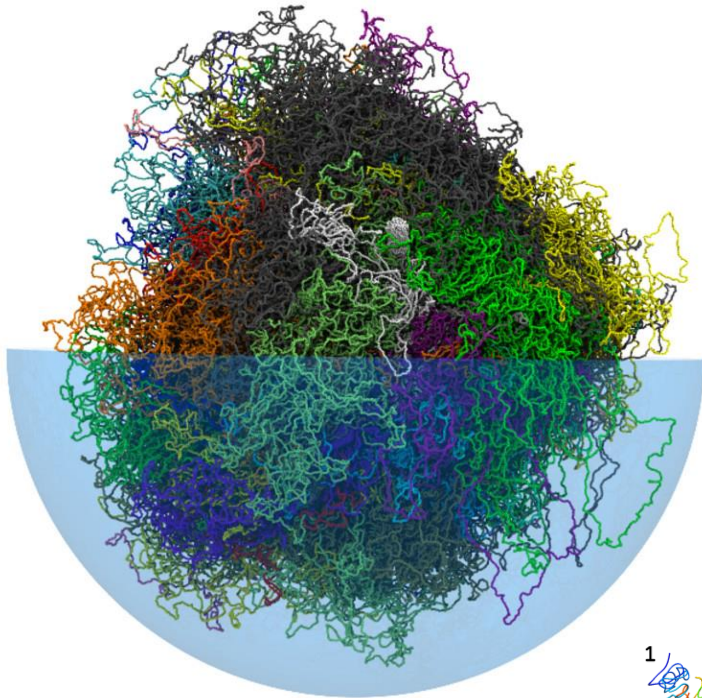
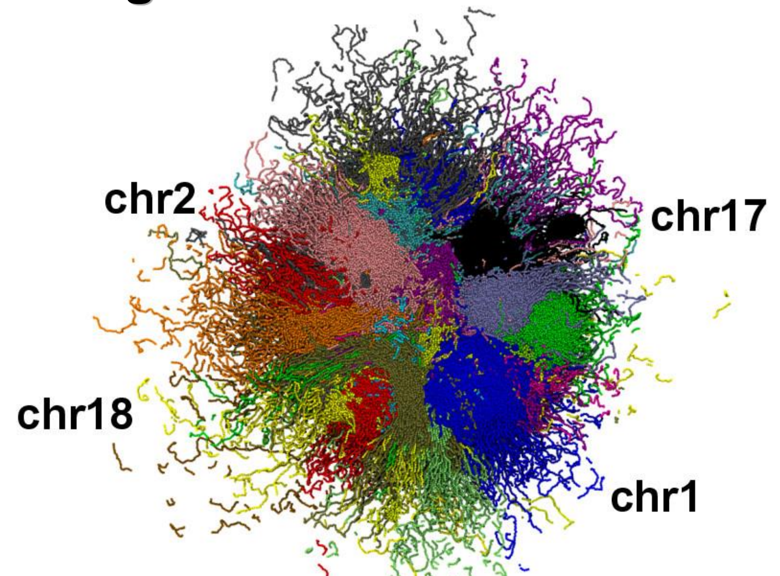
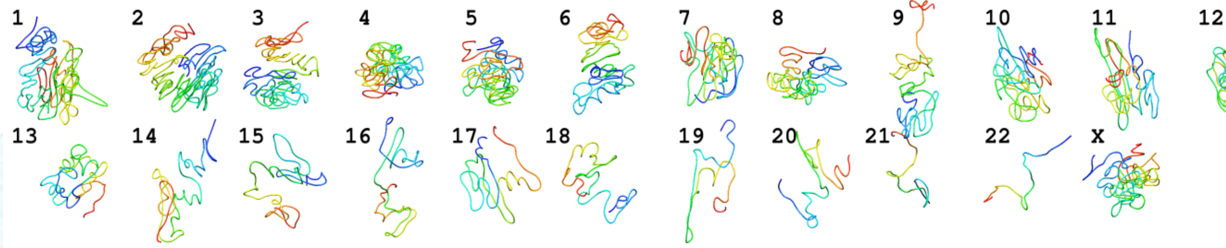
MUNI
MED

Overview of molecular biology exp. methods used in pathophysiology

Sequencing



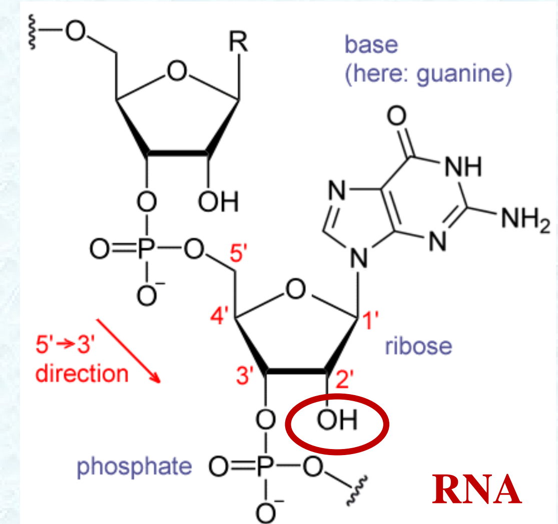


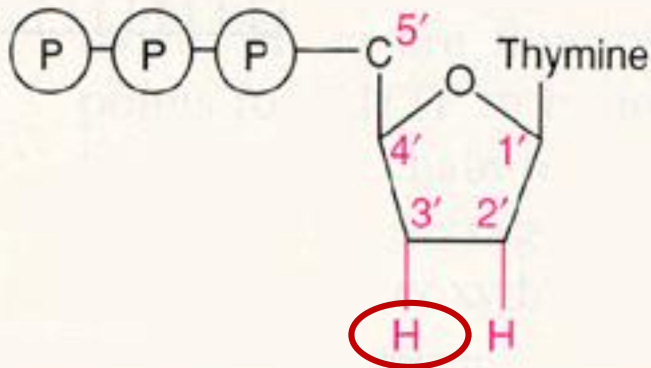


by Martin Piskacek

DNA Sequencing
developed by Frederick Sanger

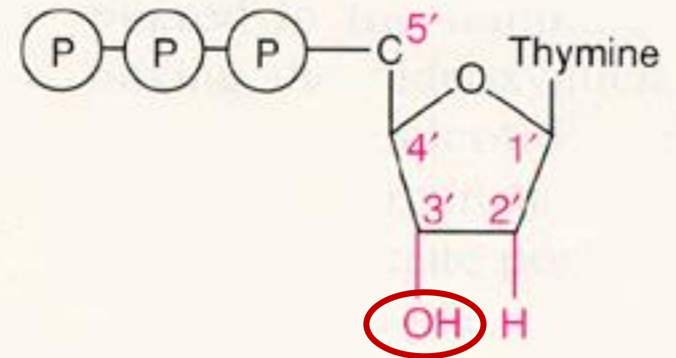
- determine the DNA sequence
- based on DNA replication
- dideoxy nucleotide missing hydroxyl group at 3' position terminate DNA extention





Dideoxythymidine triphosphate (ddTTP)

toxic nucleoside analog

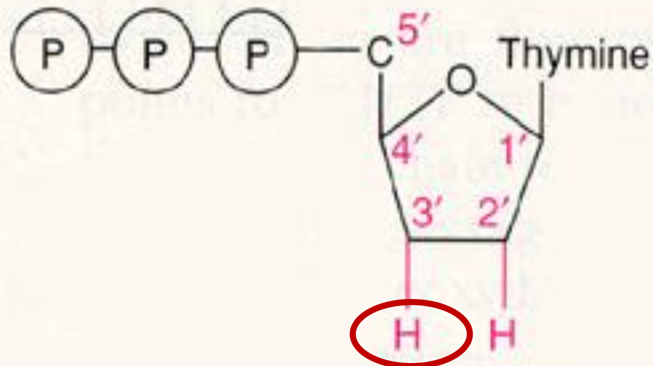


Deoxythymidine triphosphate (dTTP)

natural nucleoside in DNA

DNA synthesis and termination

- DNA chain extends in 5’→3’ direction
- Requires nucleotides 3’ –OH group
- Extension of the DNA chain is randomly terminated by dideoxynucleotides



Dideoxythymidine triphosphate (ddTTP)

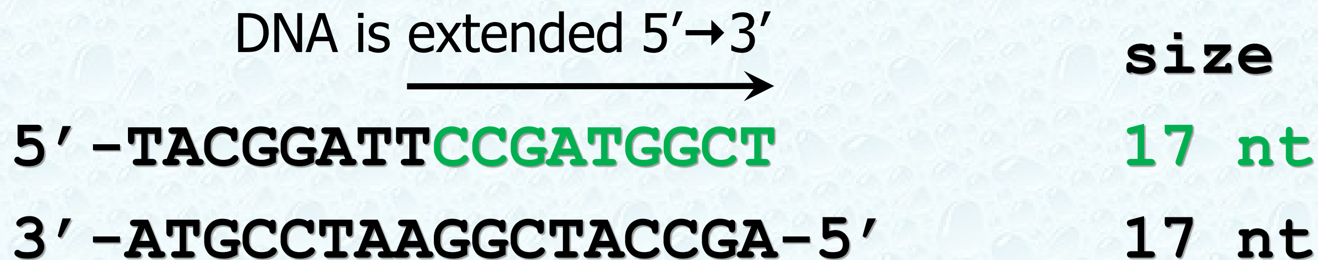
toxic nucleoside analog

Daughter strand

Template strand

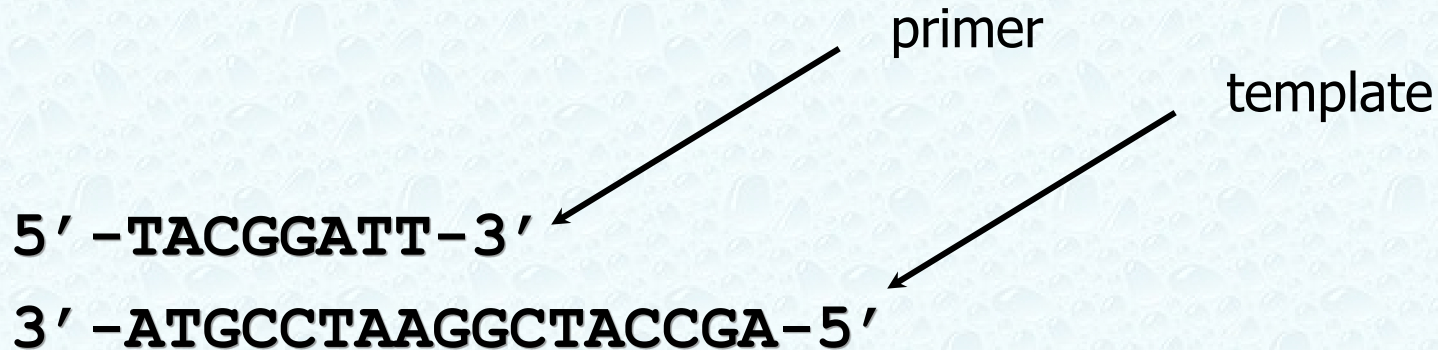
5'-to-3' direction of chain growth

DNA Extension by dATP, dGTP, dTTP, dCTP



resulted in to single band
acquired on the gel electrophoresis

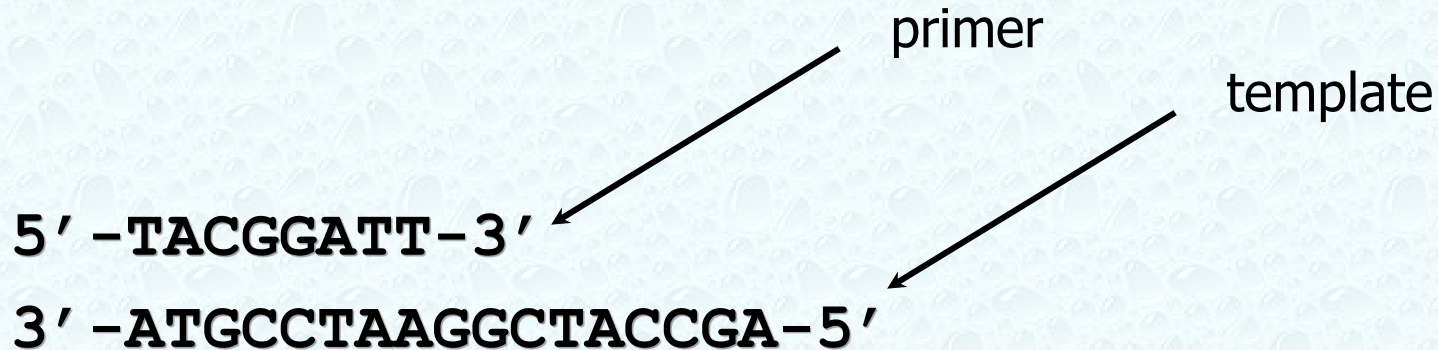
DNA Extension by dATP, dGTP, ddTTP, dCTP



size
13 nt
17 nt

resulted in to two bands
acquired on the gel electrophoresis

DNA Extension by dATP, ddGTP, dTTP, dCTP



size

11 nt

resulted in to two bands
acquired on the gel electrophoresis

DNA Extension by dATP, dGTP, dTTP, dCTP + small amount of ddATP

```
Primer:       5′-CGTCGAC-3′
Template  :   3′-GCAGCTGATGCCTATGGCTACCGA-5′
```

Reaction A	size
CGTCGACTA	9 nt
CGTCGACTACGGA	13 nt
CGTCGACTACGGATTCCGA	19 nt
CGTCGACTACGGATTCCGATGGCTA	25 nt

multiple bands
acquired on the gel electrophoresis

Dideoxy DNA Sequencing

Four separate DNA sequencing reactions are set up

Reaction G: dATP, dGTP, dTTP, dCTP + ddGTP

Reaction A: dATP, dGTP, dTTP, dCTP + ddATP

Reaction T : dATP, dGTP, dTTP, dCTP + ddTTP

Reaction C : dATP, dGTP, dTTP, dCTP + ddCTP

DNA polymerase is added to each and

- complementary strand is synthesized
- Dideoxy-nucleotides terminate synthesis

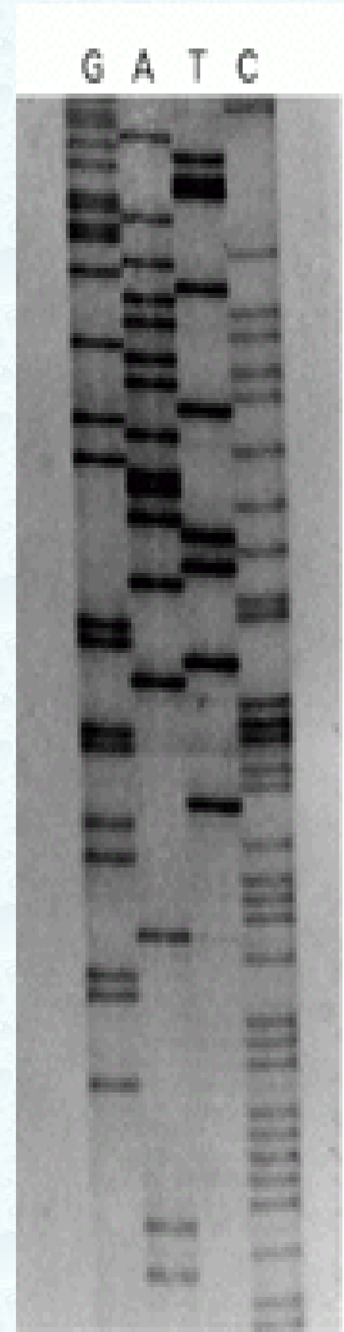
DNA fragments analytical separation

- Incorporation of a particular ddNTP produces DNA fragments
- DNA fragments are then run on a polyacrylamide gel (PAGE)
- DNA fragments, which differ by single nucleotide in size can be separated and visualized or detected
- Fragments from the four reactions (G, A, T, C) are run apart
- Location of a band on the gel corresponds to the position of that nucleotide in the linear sequence

Capillary electrophoresis

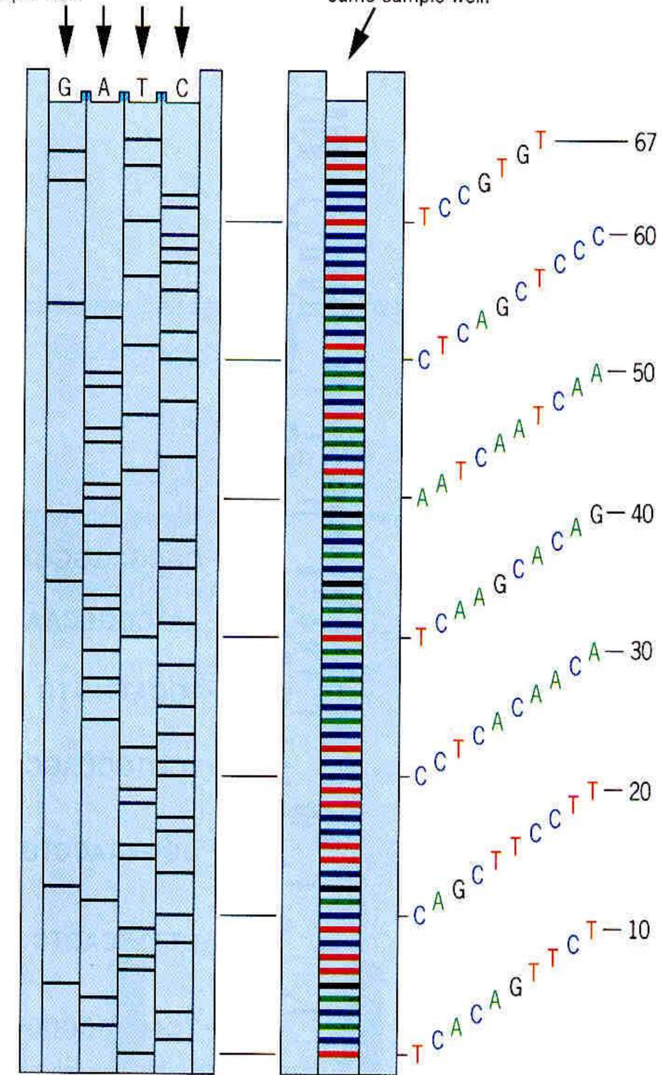
24 / 48 capillaries

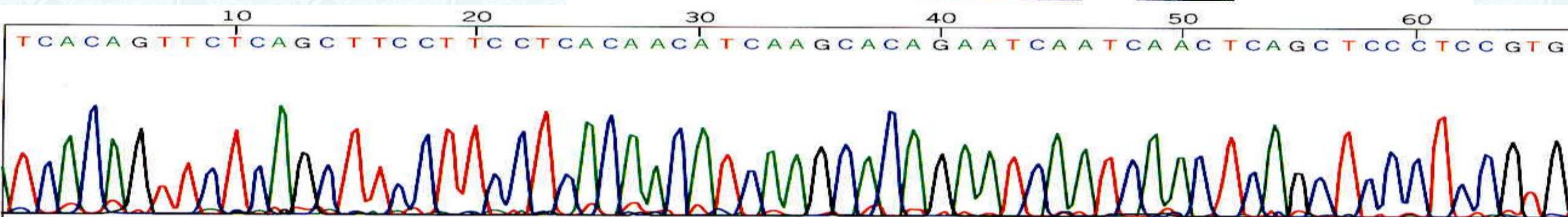
high resolution separation technique



Fluorescence Automated Sequencing

- Each of the dideoxynucleotides is labelled with a different fluorescent dye
- All four of the reactions (G, A, C, T) can take place in same tube
- The fragments are therefore loaded on to same lane of the gel
- fluorimeter and computer are linked to the gel and they detect and record the dye attached to the fragments as they come off the gel
- Sequence is determined by the order of the dyes coming off the gel



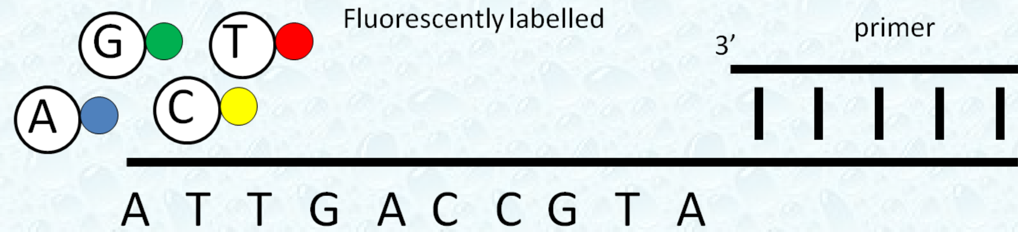


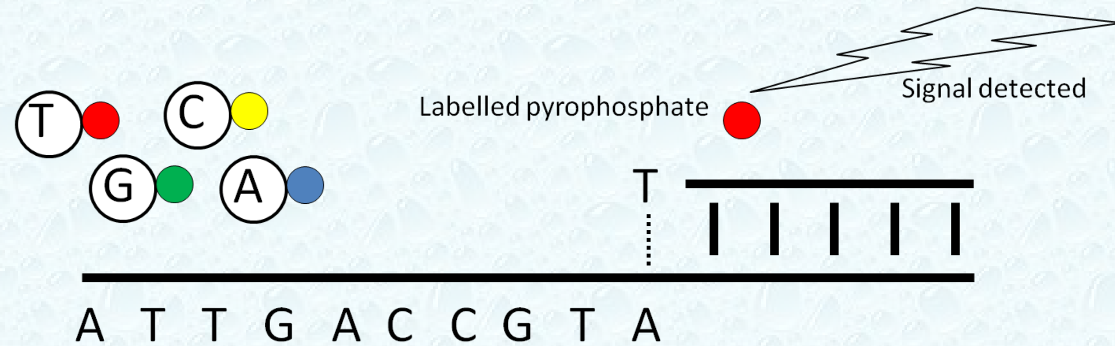
Pyrosequencing

Sequencing-by-synthesis (SBS)

- ATP acts as fuel to the luciferase-mediated conversion of luciferin to oxyluciferin that generates visible light

- Each bases has a different label on the final phosphates - label of released pyrophosphate is recorded





NANOPORE SEQUENCING

At the heart of the MinION device, an enzyme unwinds DNA, feeding one strand through a protein pore. The unique shape of each DNA base causes a characteristic disruption in electrical current, providing a readout of the underlying sequence.

DNA base

DNA double helix

Unwinding enzyme

Protein pore

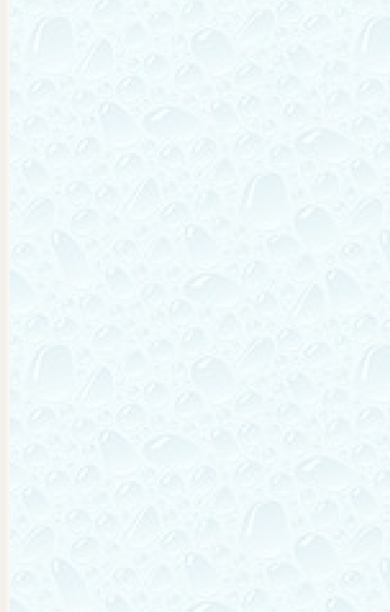
Membrane

Ion

Current

Current

Sequence A A C T C G T



Oxford Nanopore Technologies Ltd

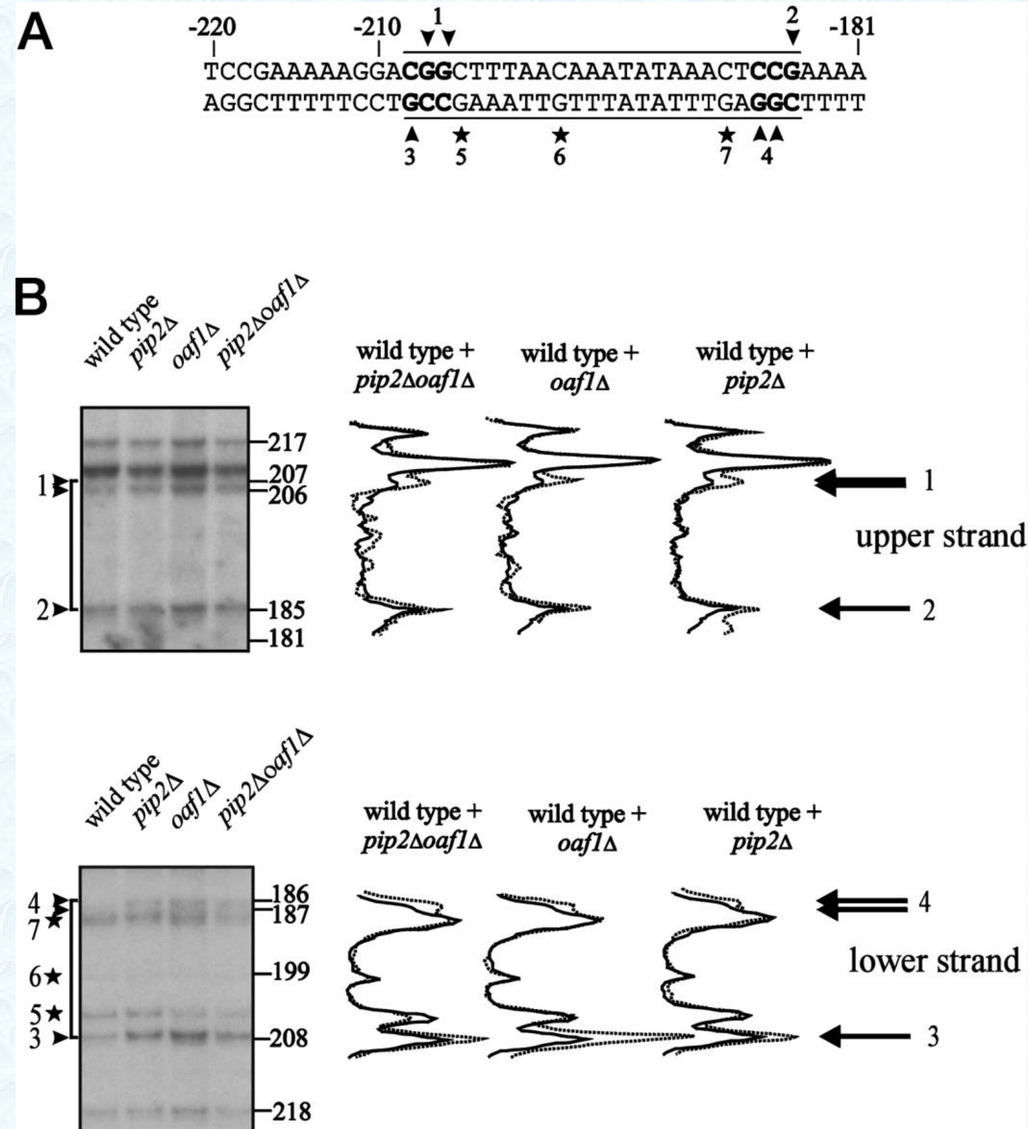
- 2048 membrane wells
- Motorprotein / Adaptor
- DNA or direct RNA seq
- 400 bp/s output monitored by current flux about 20-30 pA under 180 mV
- Adaptation of α-hemolysin α-HL nanopores for the identification of single molecules (pore-forming toxin from Staphylococcus aureus responsible for the cell lysis)
- The leader adapter guides the dsDNA fragments to the vicinity of pores, and the sequencing process begins when the leader motor protein unzips the dsDNA enabling the first strand (template) to pass through the nanopore one base at a time
- bacteriophage phi29 DNA polymerase (phi29DNAP)

Genomic DNA footprinting - applied seq.

dimethyl sulfate (DMS)
in vivo protection assays

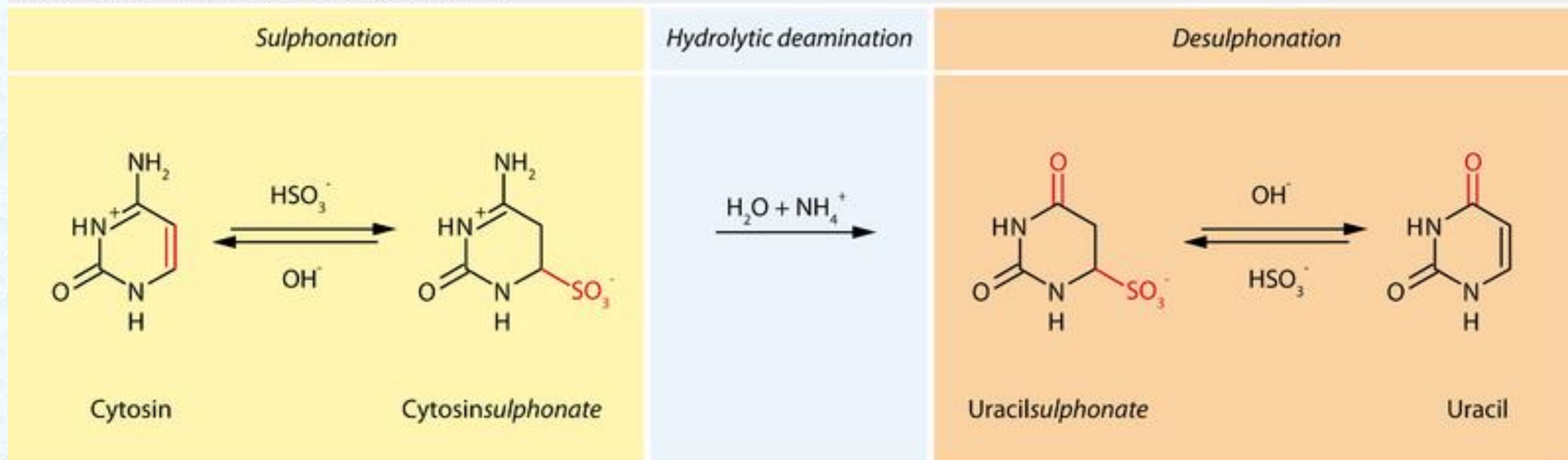
labeling / Sanger seq.
gel separation

DNAI protection
crosslinking (glutaraldehyde)
- CHIP (PCR/seq)



Methylation-specific PCR (MSP):

used to detect methylation of CpG islands in genomic DNA.
sodium bisulfite treatement , which converts
unmethylated cytosine bases in to uracil
sulphonization /de-amination /de-sulph.
MSP using qPCR can also be performed to obtain quantitative rather than qualitative information about methylation.



Allele 1 (methylated)

Allele 2 (unmethylated)

```
          m
---ACTCCACGG---TCCATCGCT---        ---ACTCCACGG---TCCATCGCT---
---TGAGGTGCC---AGGTAGCGA---        ---TGAGGTGCC---AGGTAGCGA---
                      m
```

Bisulfite treament

Alkylation

Spontaneous denaturation

```
---AUTUUAUGG---TUUATCGUT---        ---AUTUUAUGG---TUUATUGUT---

---TGAGGTGUU---AGGTAGCGA---        ---TGAGGTGUU---AGGTAGUGA---
```

Non-methylation-specific PCR

Methylation-specific PCR

Differentiation of bisulfite-generated polymorphisms

Deep Sequencing

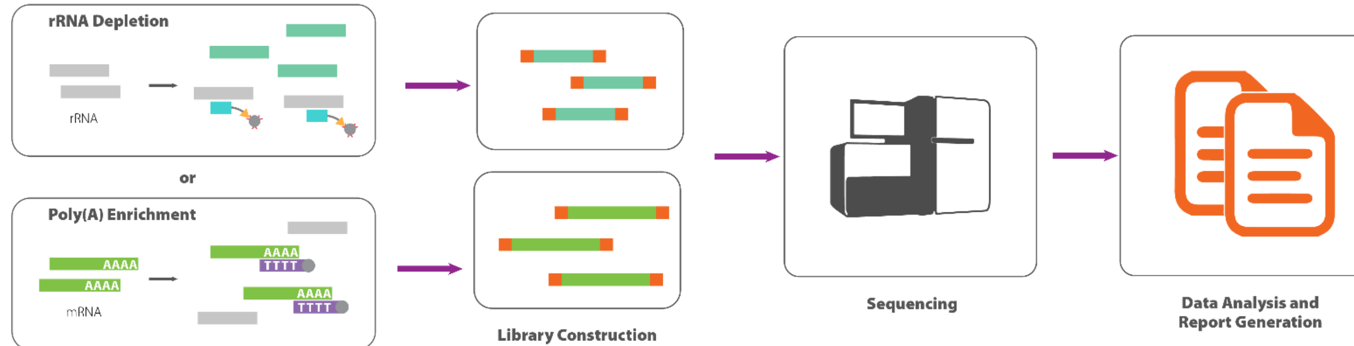
- error versus mutation
- cell clones representation
- identification of SNV

Transcriptomics

RNA-seq
alternative to **Microarray**
cDNA seq

SAGE
Serial analysis of gene expression
cDNA is generated from the RNA but is then digested into 11 bp "tag" fragments using restriction enzymes

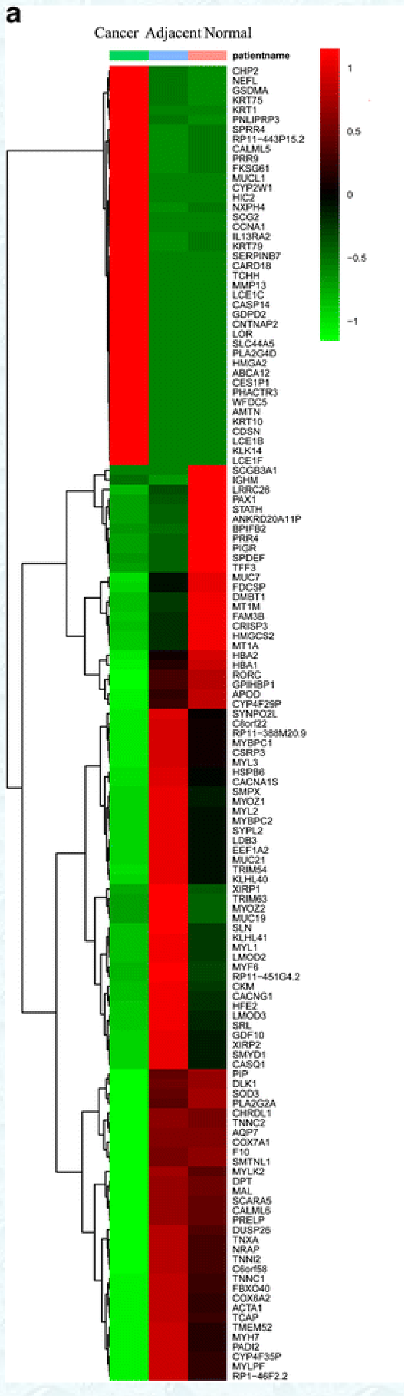
CAGE
Cap Analysis of Gene Expression
biotinylation of the 7-methylguanosine cap of Pol II transcripts, to pull down the 5'-complete cDNAs reversely transcribed from the captured transcripts



Key features for abm's RNA-Seq service	
Sequencing Platform	Illumina
Sequencing Scale	8 million reads for rapid expression analysis 40 million reads for detecting alternative splice forms 80 million reads for identifying low-abundance coding and non-coding transcripts Higher coverage available up to 400 million reads per sample
Starting Material	0.1µg – 4µg of Total RNA 10-100ng of Poly-A enriched mRNA 10-100ng of rRNA depleted RNA
Sequencing Type	75 bp single end or paired end sequencing. Longer read lengths available upon custom request.
Bioinformatics Analyses	FastQC on raw sequencing data (included) Read alignment and estimation of gene expression (included) Differential gene expression analysis Functional annotation

Project	No.	Part Number/ Description	Qty.	Unit Price (USD)	Line Total (USD)
16 human T cell of RNA library prep & sequencing **Hiseq, PE150** Q30⩾80%	1	Total RNA Isolation	16	30	480
	2	RNA library prep (rRNA depletion by Ribo-ZeroTM& directional library)	16	280	4480
	3	Sequencing of PE150 (40M reads=12G/sample)	16	240	3840
	4	Data Delivery via 1T hard disk drive	1	150	150
		Total	USD8950.00		

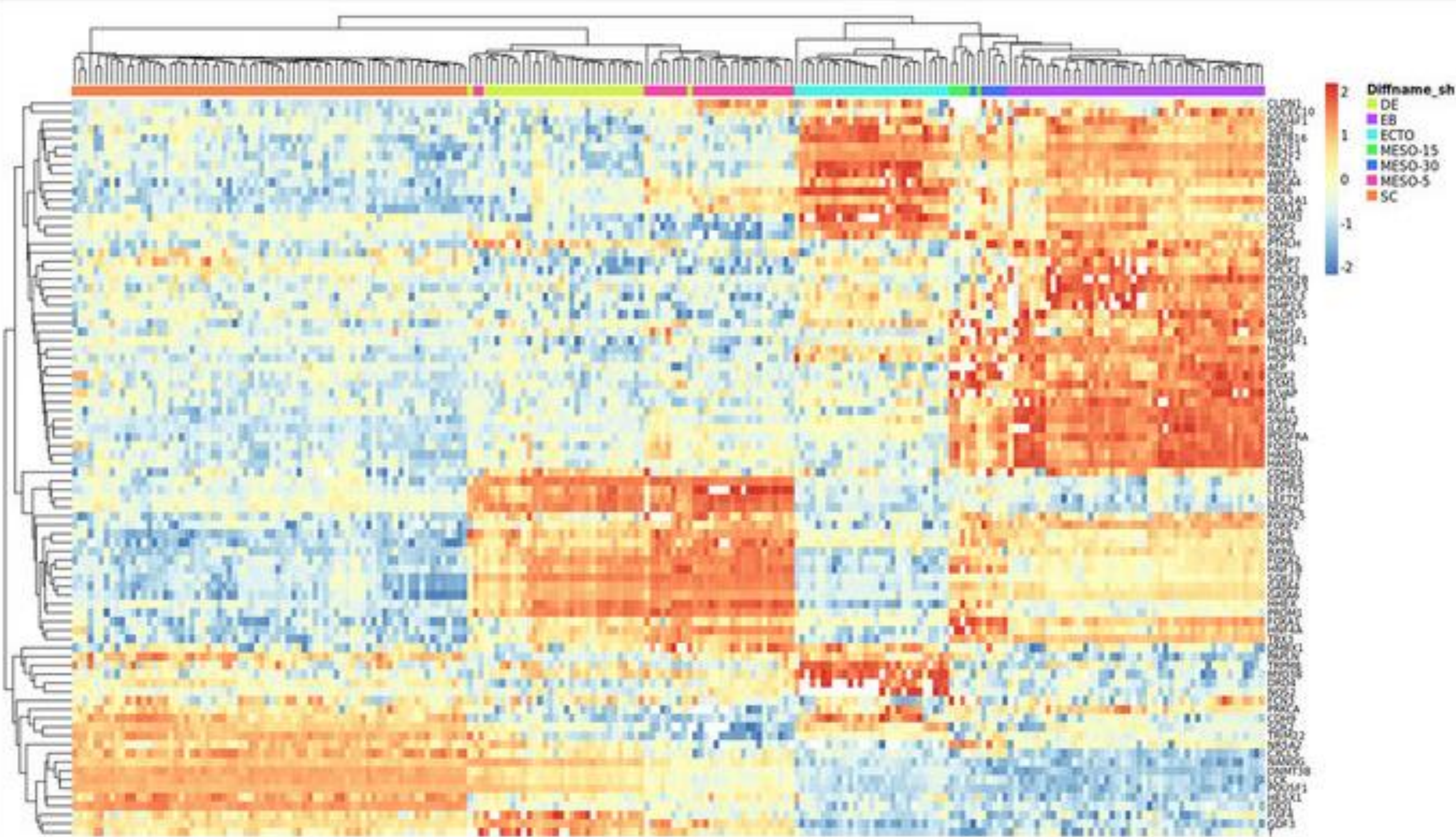
a



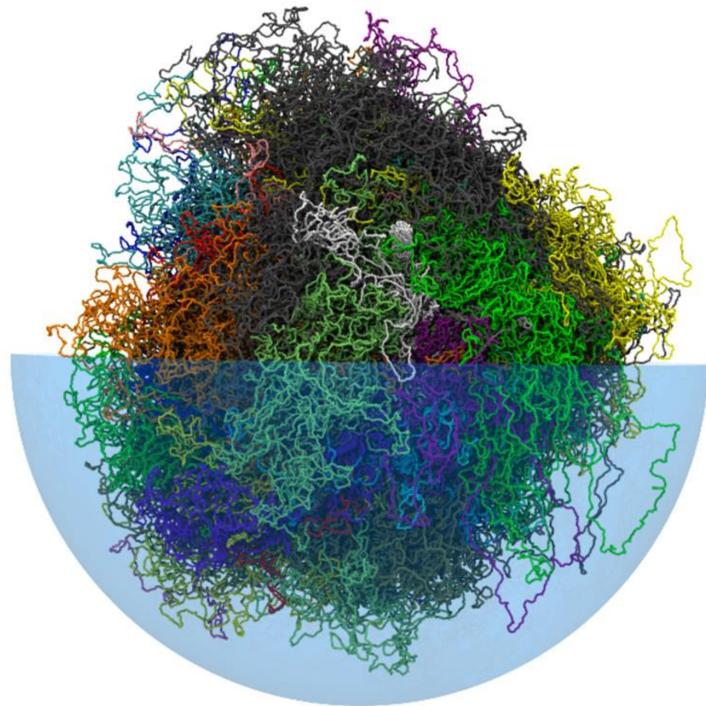
Tx
RFI
14 Tx
C
HC
12
10 RFI
High
Low
8
6
4
2

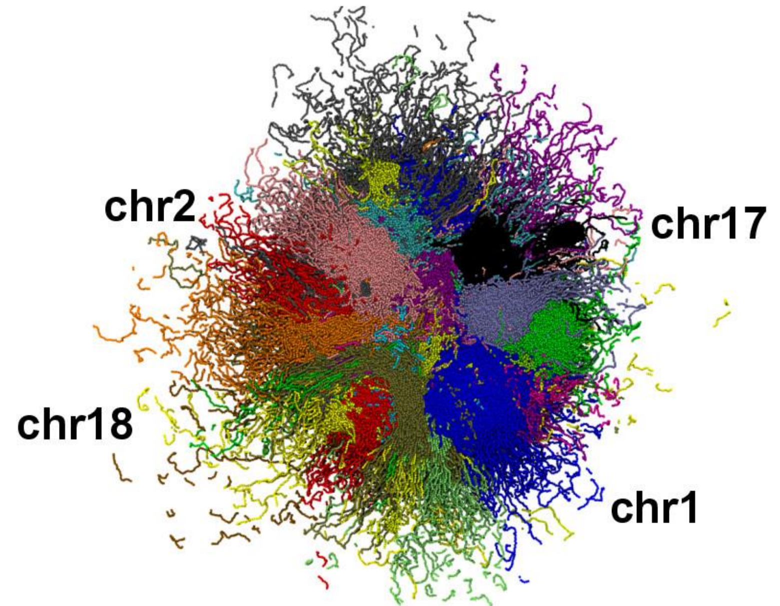
CEACAM19
CHRNE
LYZ
ARMC4
BDH2
ENSBTAG00000020499
SLC20A1
ENSBTAG00000025258
CR2
ENSBTAG00000047529
CTH
TMEM98
RYR1
CDKL4
FAM102A
C15orf48
UHRF1
CYP11A1
ARHGEF37
CYP7A1
MYOM1
CCDC64
SLC25A34
ELOVL6
SLC17A1
OXER1
ENSBTAG00000000170
LAMB3
FOSL2
CCDC109B
INHBE
SOCS2
ENSBTAG00000020375
ENSBTAG00000046076
ACOT8
KIAA1462
CD3D
ACACA
FGFR2
SELP
MTFP1
CKB
ANXA5
ENSBTAG00000010463
NKG7
IFI6
SLC6A11
DGAT2
ENSBTAG00000046730
REC8
FICD
MBOAT2
ABR
JSP.1
SLA
HCLS1
HACL1
ALDH18A1
HSD17B4
AIM1L
SLC35A5
TBC1D8
TAF6
RND3
NR1H4
MFSD1
INTS7
ENSBTAG00000037913
PCSK6
SEMA4G

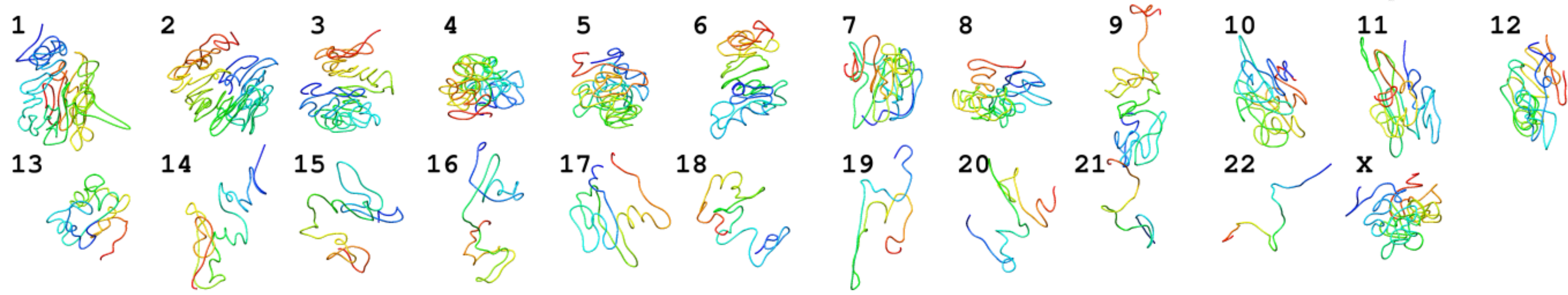
6118_HC
6118_C
5957_C
5957_HC
5682_C
5682_HC
5790_C
5790_HC
6199_C
6199_HC
6144_C
6144_HC
6005_C
6005_HC
5751_C
5751_HC
5544_C
5544_HC

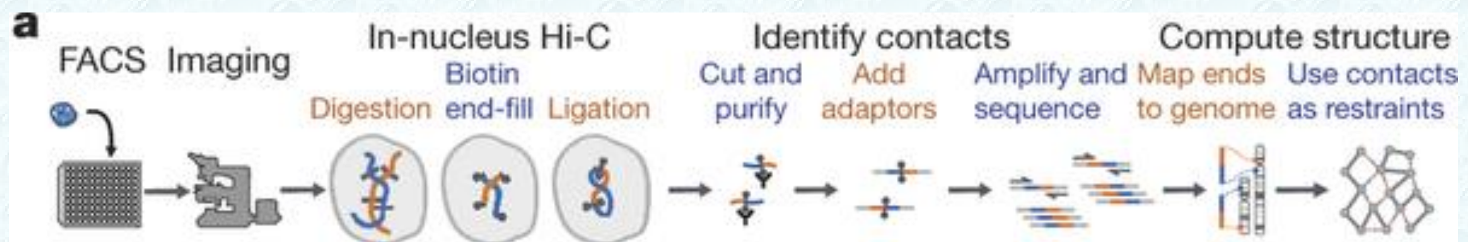

Diffname_sh
DE
EB
ECTO
MESO-15
MESO-30
MESO-5
SC
2
1
0
-1
-2

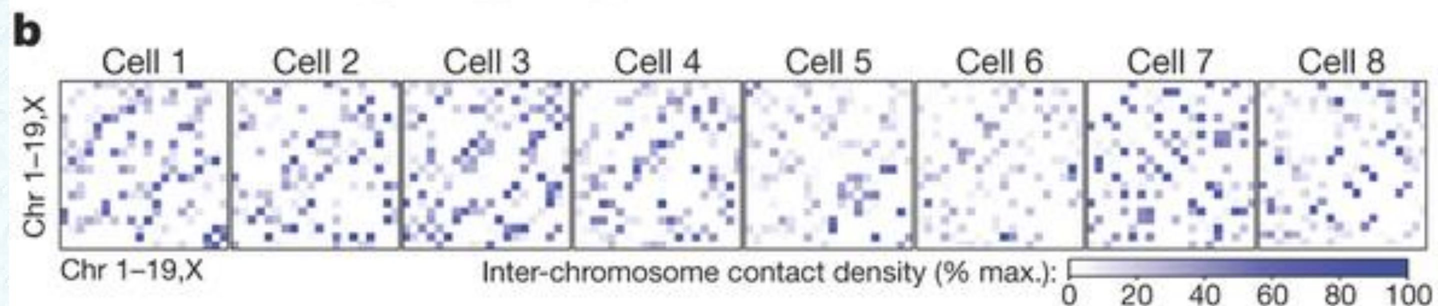
Hi-C all-vs-all
high-throughput chromosome conformation capture
cross linking + fragment ligation
high throughput fragment sequencing











c

Chr 10

Nuclear position Chr 10

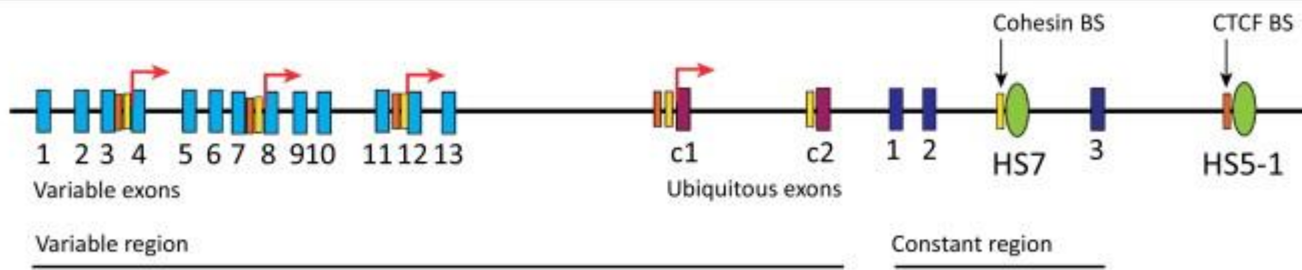
Cell 1 – all chromosomes

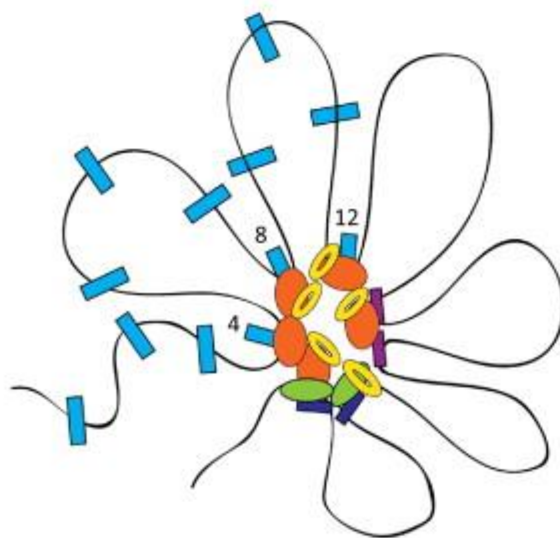
Restraints Chr 10

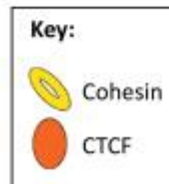
(5 models)

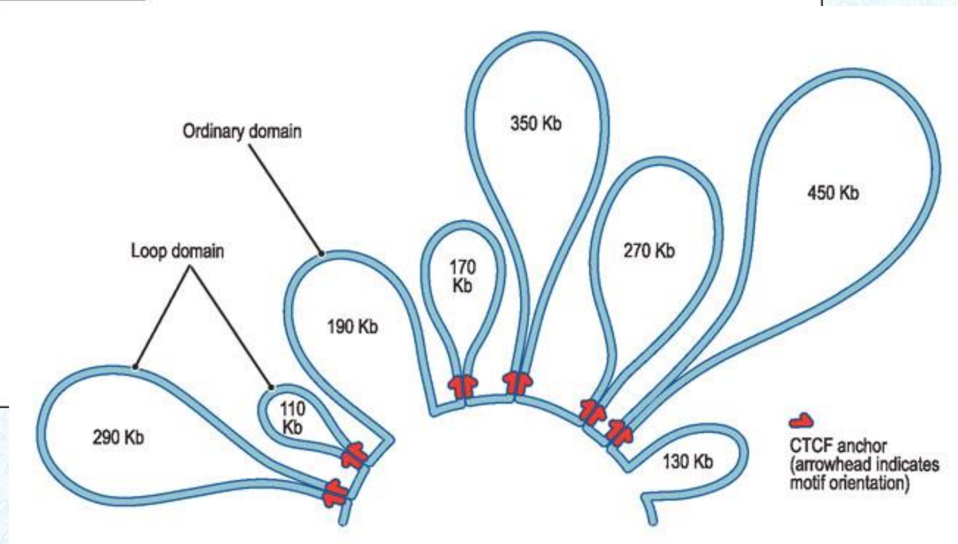
(5 models)

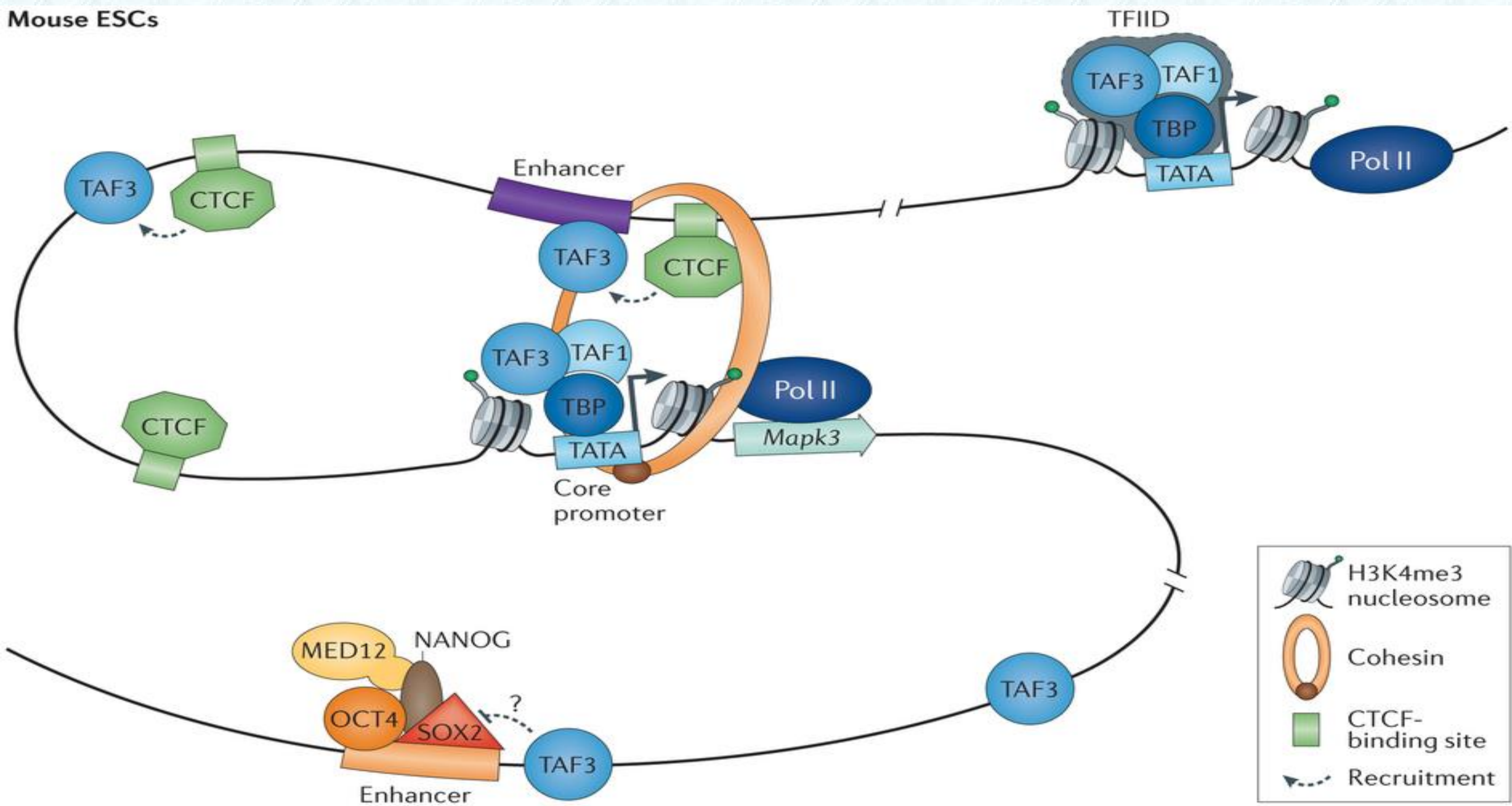
Topologically Associating Domains TADs



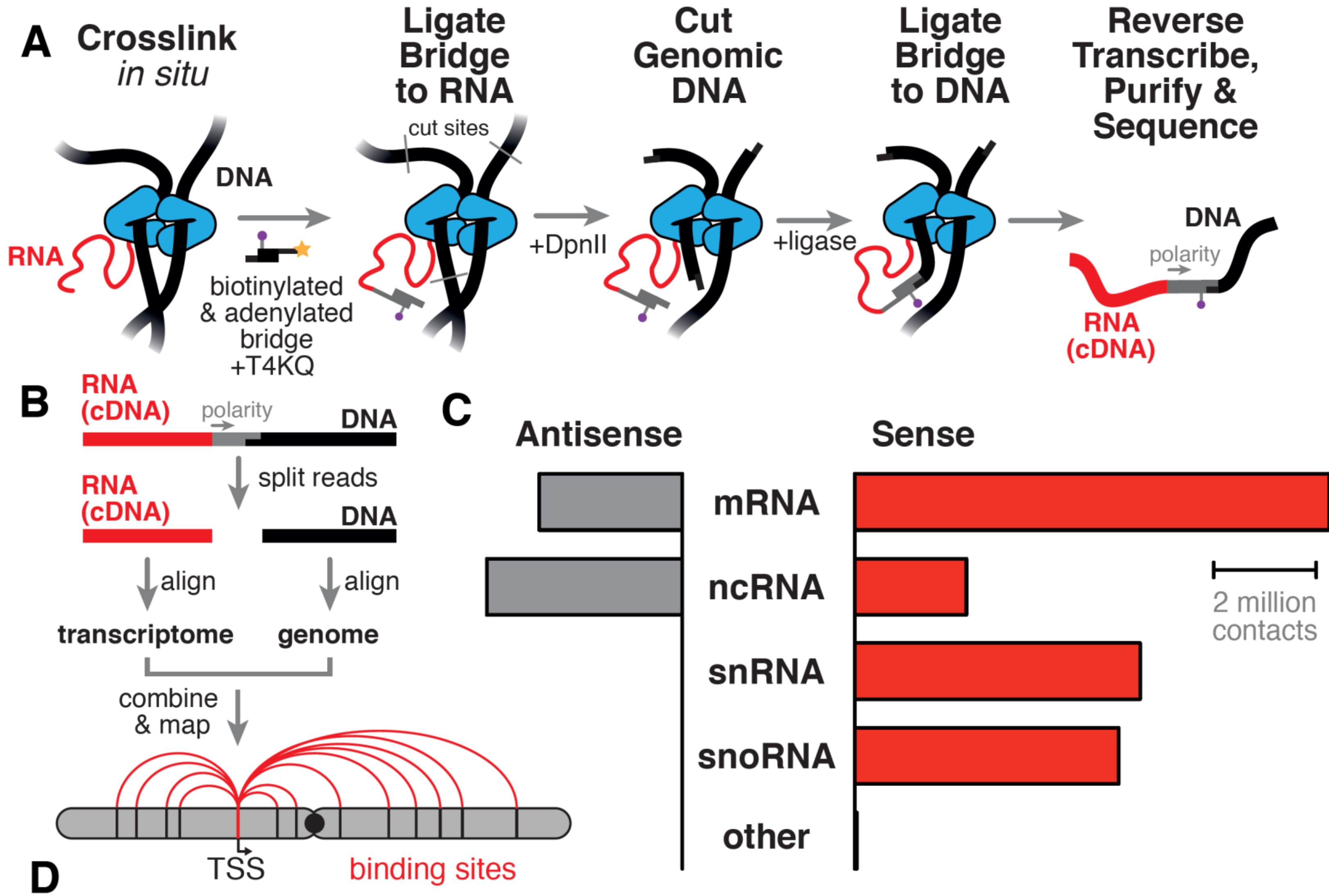




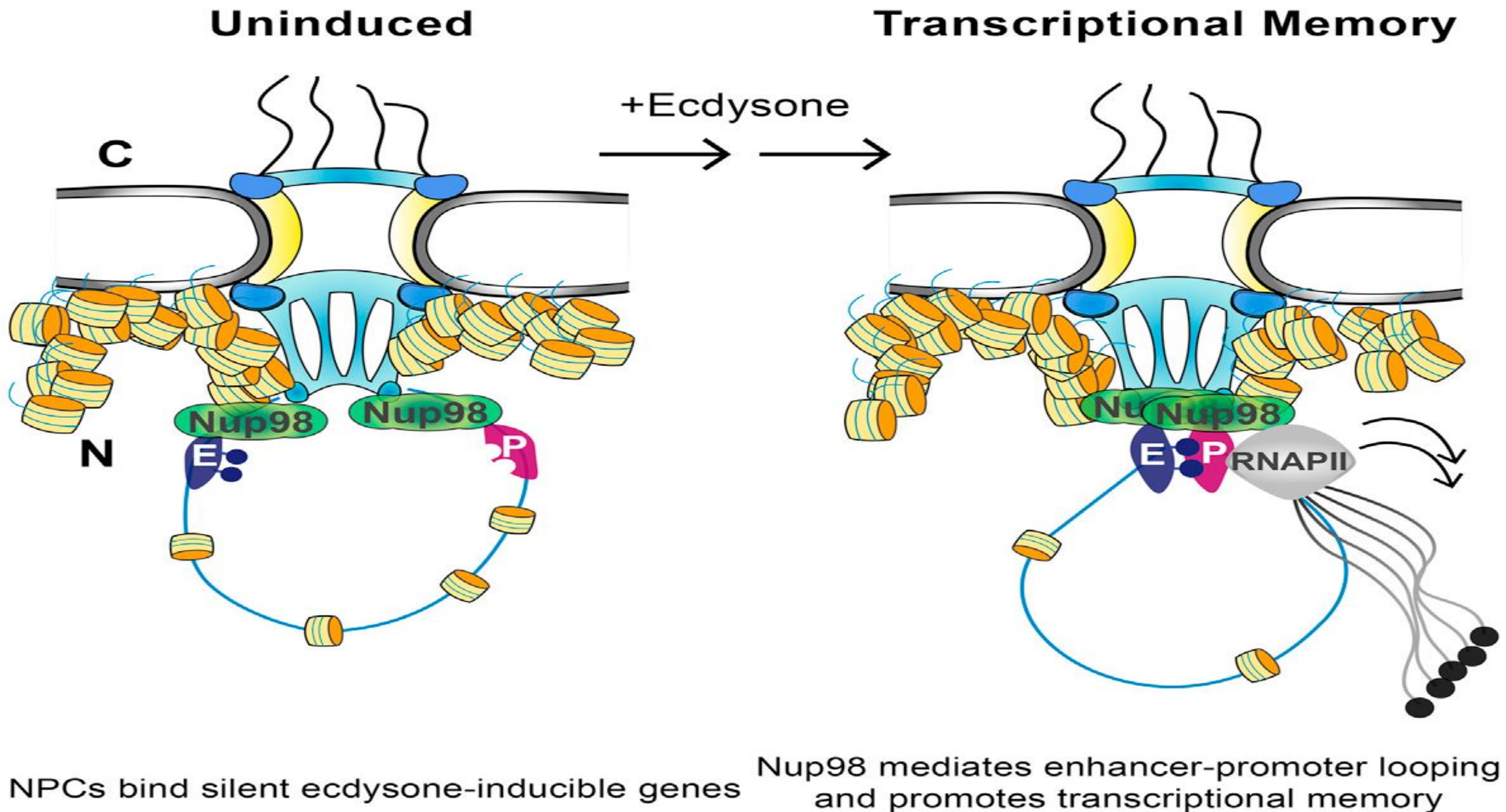



Mouse ESCs
TFIID
TAF3
TAF1
TBP
TATA
Pol II
TAF3
CTCF
Enhancer
TAF3
CTCF
TAF3
TAF1
TBP
TATA
Pol II
Mapk3
CTCF
Core promoter
MED12
NANOG
OCT4
SOX2
?
TAF3
Enhancer
TAF3
H3K4me3 nucleosome
Cohesin
CTCF-binding site
Recruitment
Nature Reviews | Genetics

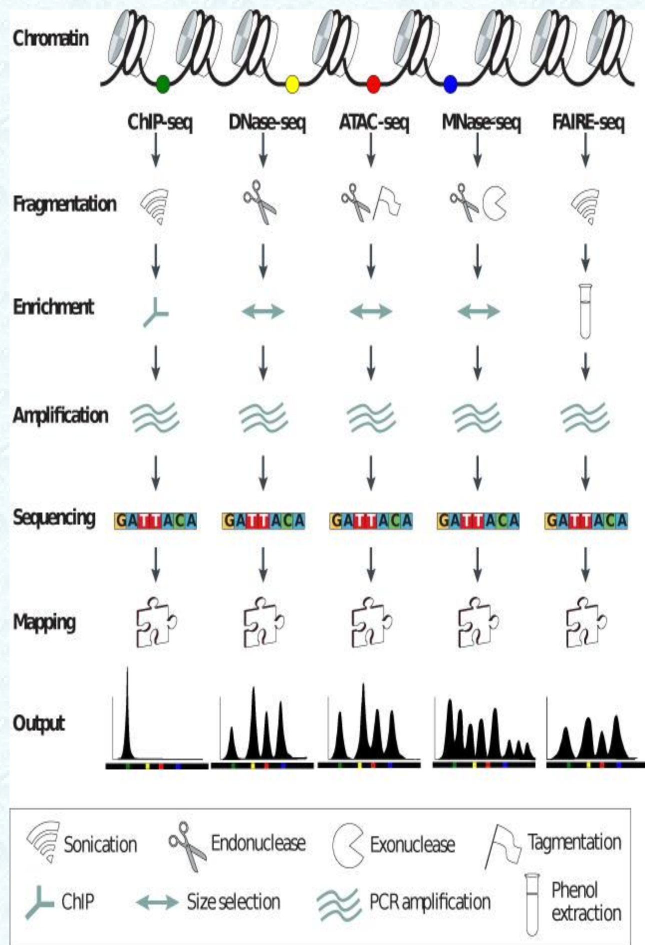
Topologically Associating Domains TADs and chromatin associated RNA



Nuclear Pores Bind Inducible Genes
Nup98 ChIP-seq



Monitoring of nucleosome occupancy in the human genome:

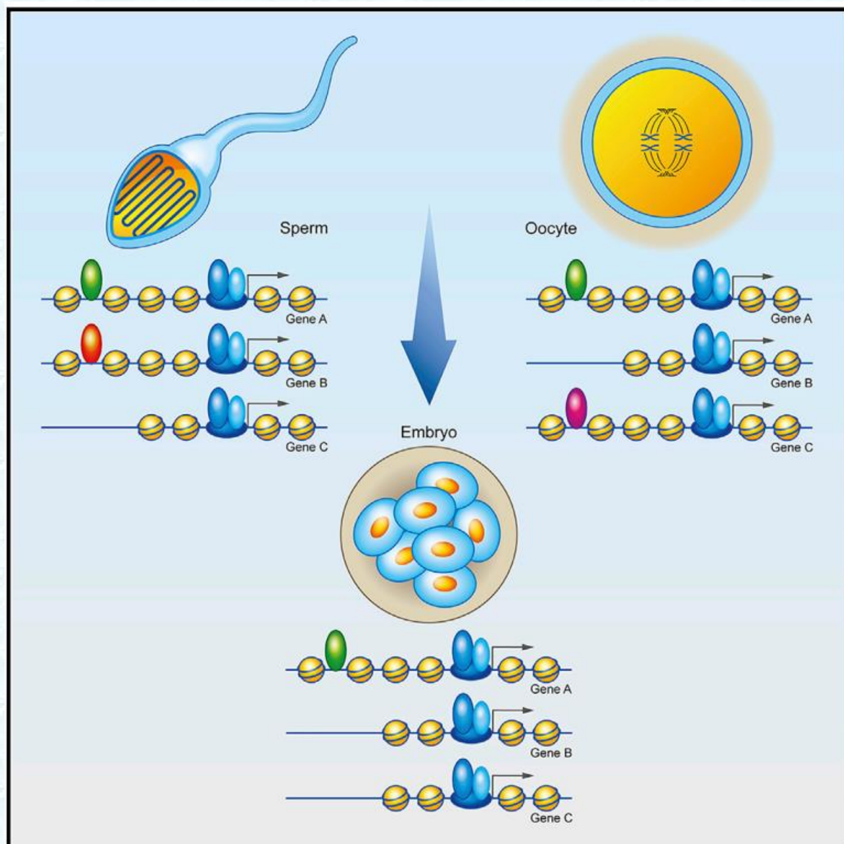


ChIP-seq specifically and genome-wide assay
chromatin immunoprecipitation with DNA sequencing, to obtain the DNA sequences which interact with transcription factors and nucleosome-associated sequences

DNase / MNase-seq chromatin accessibility assay genome-wide
DNase I hypersensitive sites sequencing
micrococcal **nuclease digestion** (Staphylococcus aureus)
identification of accessible DNA regions in the genome
DNA bound to histones or other chromatin-bound proteins
transcription factors / repressors / regulators SWI may **remain undigested**

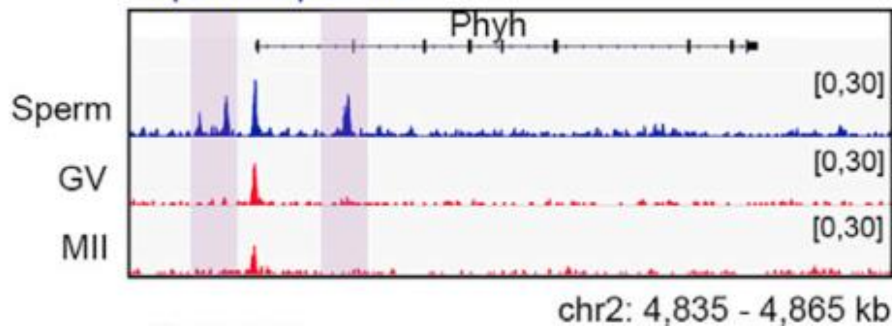
- ATAC-seq THSS - **Tn5 hypersensitive site**
Assay for Transposase-Accessible Chromatin using sequencing
mutated hyperactive transposase Tn5
transposase inserts sequencing adapters into unprotected regions

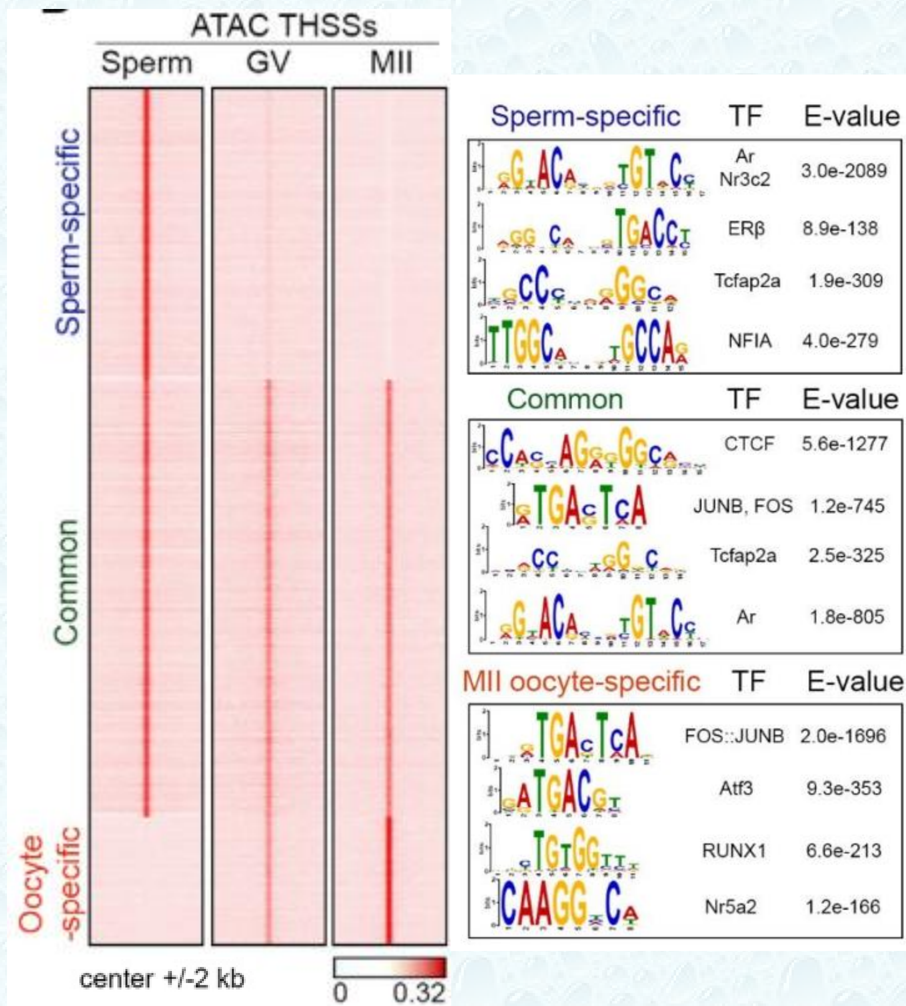
- FAIRE-seq
Formaldehyde-Assisted Isolation of Regulatory Elements
formaldehyde cross-link DNA and proteins, sonication and **phenol-chloroform extraction -**nucleosomes will preferentially sit in the organic phase and **nucleosome-depleted regions** will be purified from aqueous phase

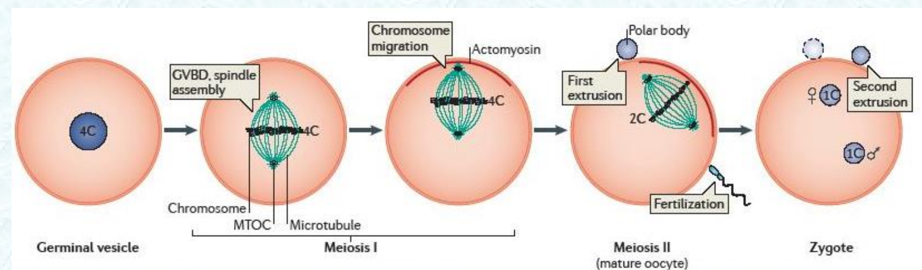


Highlights

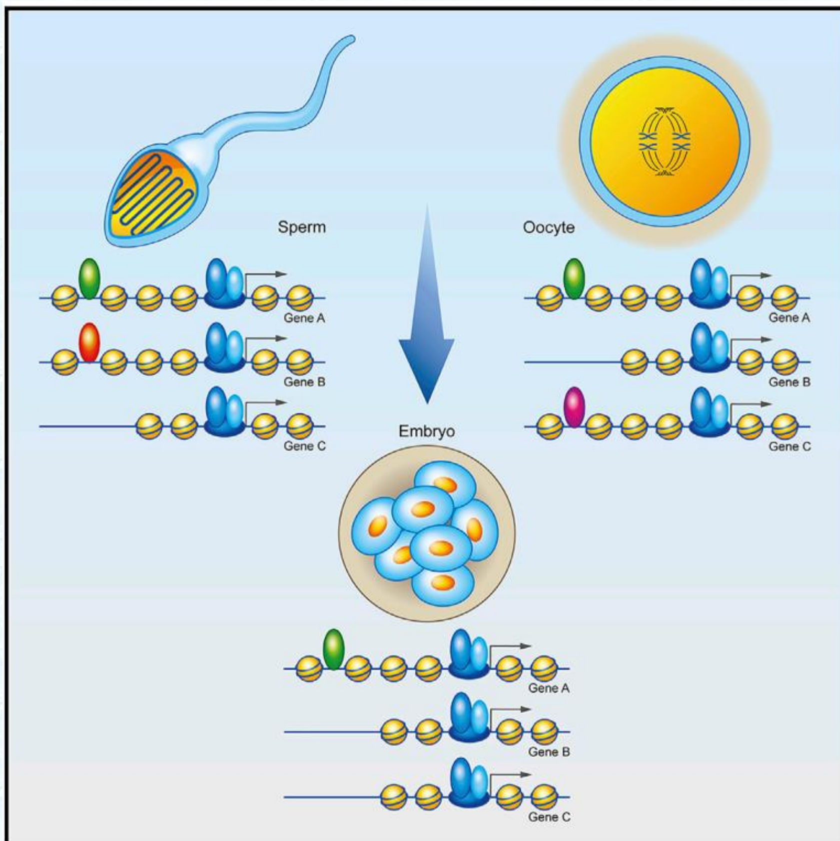
- **ATAC-seq accessibility at sperm and oocyte promoters is maintained in the embryo**







Jung et al., 2019, Molecular Cell 75, 1–18
July 11, 2019 © 2019 Elsevier Inc.
https://doi.org/10.1016/j.molcel.2019.04.014



ChIP-seq

Heatmaps and average profiles showing Foxa1, ERa, and AR ChIP-seq signal co-localized around transcription start sites TSS in sperm chromatin at all Refseq annotated genes.
Sites are ordered by k-means clustering of RNAPIISer5ph and RNAPIISer2ph signal between TSSs and TTSs