Background
Characterizing nucleosome dynamics
Conclusions and future works

Characterizing nucleosome dynamics from genomic and epigenetic information using rule induction learning

Ngoc Tu Le   Tu Bao Ho   Dang Hung Tran

Japan Advanced Institute of Science and Technology

Singapore $*$ September 9-11, 2009

Background
Characterizing nucleosome dynamics
Conclusions and future works

Outline

Background
Basic concepts
Related works

Characterizing nucleosome dynamics
Method overview
Results

Conclusions and future works
Conclusions
Tentative works

Background
Characterizing nucleosome dynamics
Conclusions and future works

Basic concepts
Related works

- ▶ The ability of cells in living organisms depends on many biological processes
    - ▶ *RNA and protein synthesis*
    - ▶ *Replication*
    - ▶ *DNA repair* and
    - ▶ *Genetic recombination*
- ▶ Biological machineries, such as *Transcription Initiation Complex* (TIC), etc
- ▶ Regulatory regions, such as *promoters*, *enhancers*, *silencers*, etc

Background
Characterizing nucleosome dynamics
Conclusions and future works

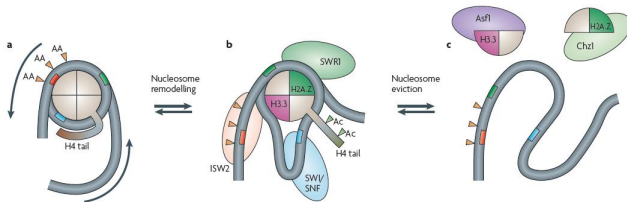Basic concepts
Related works
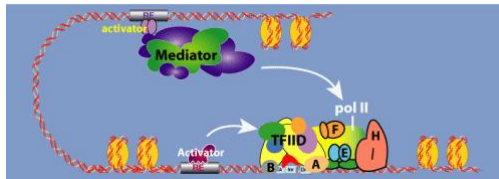
Chromatin structure
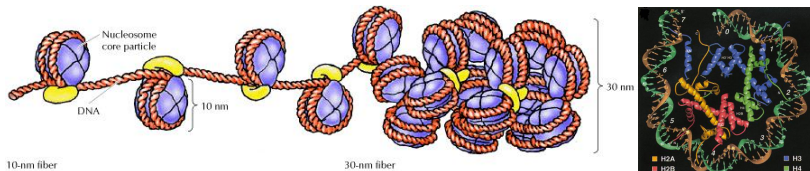
- ▶ Eukaryotic genome is organized under chromatin structure containing basic units of *nucleosomes*
- ▶ Each nucleosome is a complex of 147bp DNA coiling around a histone core of protein octamer, one H3-H4 tetramer and two H2A-H2B dimers

Background
Characterizing nucleosome dynamics
Conclusions and future works

Basic concepts
Related works

Destabilizing nucleosomes - A way for gaining access to nucleosomal DNA

-Jiang C and Pugh BF, Nat Rev Genet. 2009-

▶ Biological machineries *in vivo* gain access to DNA by dislocating or even evicting nucleosomes [Li B et al., 2007; Jiang C and Pugh BF, 2009]

▶ Promoters of actively expressed genes are usually free of nucleosomes [Lee CK et al., 2004; Henikoff S, 2007]

Background
Characterizing nucleosome dynamics
Conclusions and future works

Basic concepts
Related works

Problems

- ▶ What factors affect nucleosome dynamics? and How?
- ▶ What is the relationship between nucleosome dynamics and gene regulation?

Background
Characterizing nucleosome dynamics
Conclusions and future works

Basic concepts
Related works

Problems

- ▶ What factors affect nucleosome dynamics? and How?
- ▶ What is the relationship between nucleosome dynamics and gene regulation?

Background
Characterizing nucleosome dynamics
Conclusions and future works

Basic concepts
Related works

- ▶ Nucleosome dynamics can be affected by many factors [Henikoff S, 2008; Segal E and Widom J, 2009; Marta RL and Rando OJ, 2009]
    - ▶ DNA sequences
    - ▶ Histone modifications and histone variants
    - ▶ Chromatin remodelling complexes and histone chaperones
- ▶ Experiment data on model organisms
    - ▶ Nucleosome positioning [Yuan GC et al., 2005; Lee W et al., 2007; Mavrich TN et al., 2008]
    - ▶ Histone modifications [Pokholok DK et al., 2005; Liu CL et al., 2005; Barski A et al., 2007]
- ▶ Previous works have some common drawbacks
    - ▶ Didn't concern much about destabilized nucleosomes
    - ▶ Investigated the effect of each factor individually
    - ▶ Based mainly on experiments

Background
Characterizing nucleosome dynamics
Conclusions and future works

Basic concepts
Related works

- ▶ Nucleosome dynamics can be affected by many factors [Henikoff S, 2008; Segal E and Widom J, 2009; Marta RL and Rando OJ, 2009]
    - ▶ DNA sequences
    - ▶ Histone modifications and histone variants
    - ▶ Chromatin remodelling complexes and histone chaperones
- ▶ Experiment data on model organisms
    - ▶ Nucleosome positioning [Yuan GC et al., 2005; Lee W et al., 2007; Mavrich TN et al., 2008]
    - ▶ Histone modifications [Pokholok DK et al., 2005; Liu CL et al., 2005; Barski A et al., 2007]
- ▶ Previous works have some common drawbacks
    - ▶ Didn't concern much about destabilized nucleosomes
    - ▶ Investigated the effect of each factor individually
    - ▶ Based mainly on experiments

Background
Characterizing nucleosome dynamics
Conclusions and future works

Method overview
Results

- ▶ Identify significant DNA motifs by two approaches
  - ▶ Motif finding
  - ▶ Feature selection
- ▶ Identify significant histone modifications by feature selection
- ▶ Induce rules characterizing nuclesome dynamics
- ▶ Data
  - ▶ Nucleosomal sequences and nucleosome states [Yuan GC et al., 2005]
  - ▶ 12 kinds of histone modifications [Liu CL et al., 2005]
  - ▶ 2 datasets, chromosome III and promoter regions

Background
Characterizing nucleosome dynamics
Conclusions and future works

Method overview
Results
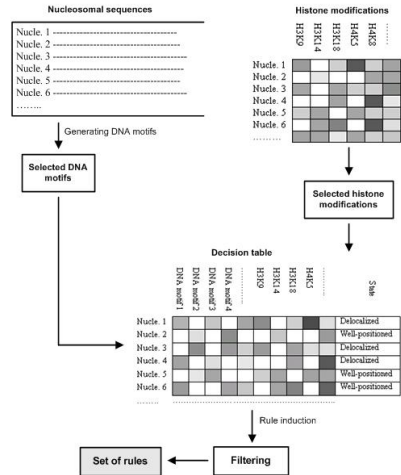
Significant motifs to nucleosome dynamics

- ▶ Motifs given by *WordSpy* [Wang G et al., 2006] with lengths from 2 to 6
  - ▶ Statistically significant motifs
  - ▶ Not much difference between chromosome III and promoter regions
- ▶ Discriminative motifs ranked by *Fisher criterion* (F-score) given by *GIST* [Pavlidis P et al., 2004]
  - ▶ Dinucleotide is the strongest discriminative signal
  - ▶ Several are related with nucleosome forming (CA/TG/CC/GC) and inhibiting (AT/TA/TT) [Peckham HE et al., 2007]

Background
Characterizing nucleosome dynamics
Conclusions and future works

Method overview
Results

Significant histone modifications to nucleosome dynamics

- ▶ Histone modifications ranked by F-score
- ▶ On chromosome III: 9 significant modifications
- ▶ On promoter regions: 6 significant modifications
- ▶ Modifications related with transcriptional activities: H3K4Me3/H3K9Ac/H3K18Ac (activation) and H4K16Ac/H4K12Ac/H4K8Ac (repression) [Kurdistani SK et al., 2004; Liu CL et al., 2005; Barski A et al., 2007]

Background
Characterizing nucleosome dynamics
Conclusions and future works

Method overview
Results

Effects on nucleosome dynamics

No.	Rules
1	$AA, ATT = enr \land H3K9Ac = neutral \rightarrow State = Well$
2	$ATT = enr \land H3K4Me3 = hyper \rightarrow State = Well$
3	$AT, GC = enr \land CC = low \rightarrow State = Well$
4	$AT, CC = enr \land GC = low \rightarrow State = Well$
5	$AT = low \land H3K9Ac = neutral \land H4K12Ac = hyper \rightarrow State = Well$
6	$AT, TC = low \land ATT = enr \rightarrow State = Well$
7	$CT, TG, GA, AT, CTT, GAG, ATT = low \land H3K18Ac, H3K4Me3 = hyper \rightarrow State = Del$
8	$GA, TT, GG = low \land H3K9Ac = hyper \land H3K4Me3 = hypo \rightarrow State = Del$
9	$AA = low \land GT, ATT = enr \rightarrow State = Well$
10	$ATT = enr \land H3K9Ac = hyper \rightarrow State = Well$
11	$GA, AG, ATT = low \land H2BK16Ac = neutral \land H4K12Ac = hypo \rightarrow State = Del$
12	$AT = enr \land TA, TAA = low \land H3K9Ac = neutral \land H4K12Ac = hypo \rightarrow State = Del$

- ▶ Rules induced by *CN2-SD* [Lavrac N et al., 2004]: 60 on chromosome III and 38 on promoter regions
- ▶ The enrichment of some motifs (e.g AT/ATT/CTT) has strong effect (rules 1,2,6,9,10)
- ▶ Some modifications (e.g H3 K9Ac/K18Ac/K4Me3, H4K12Ac) have large impact (rules 5,7,8,11,12)

Background
Characterizing nucleosome dynamics
Conclusions and future works

Method overview
Results

Effects on nucleosome dynamics

No.	Rules
1	$AA, ATT = enr \land H3K9Ac = neutral \rightarrow State = Well$
2	$ATT = enr \land H3K4Me3 = hyper \rightarrow State = Well$
3	$AT, GC = enr \land CC = low \rightarrow State = Well$
4	$AT, CC = enr \land GC = low \rightarrow State = Well$
5	$AT = low \land H3K9Ac = neutral \land H4K12Ac = hyper \rightarrow State = Well$
6	$AT, TC = low \land ATT = enr \rightarrow State = Well$
7	$CT, TG, GA, AT, CTT, GAG, ATT = low \land H3K18Ac, H3K4Me3 = hyper \rightarrow State = Del$
8	$GA, TT, GG = low \land H3K9Ac = hyper \land H3K4Me3 = hypo \rightarrow State = Del$
9	$AA = low \land GT, ATT = enr \rightarrow State = Well$
10	$ATT = enr \land H3K9Ac = hyper \rightarrow State = Well$
11	$GA, AG, ATT = low \land H2BK16Ac = neutral \land H4K12Ac = hypo \rightarrow State = Del$
12	$AT = enr \land TA, TAA = low \land H3K9Ac = neutral \land H4K12Ac = hypo \rightarrow State = Del$

▶ No DNA motif or modification shows dominant effect
▶ There exist combinatorial effects
  ▶ by DNA motifs themselves (rules 3,4,9) [Peckham HE et al., 2007; Jiang C and Pugh BF, 2009] or
  ▶ by both DNA motifs and histone modifications (rules 2,5,7,8,10,11,12) [Wang A et al., 2002; de Nadal E et al., 2004]

Background
Characterizing nucleosome dynamics
Conclusions and future works

Conclusions
Tentative works

- ▶ Proposed method is quite suitable for characterizing nucleosome dynamics
- ▶ It can discover potential co-effects of DNA sequences and histone modifications on nucleosome dynamics
    - ▶ Some DNA motifs and histone modifications are more important to nucleosome dynamics
    - ▶ They not only act individually but also cooperate with each other to regulate nucleosome stability
- ▶ Other factors also need to be considered

Background
Characterizing nucleosome dynamics
Conclusions and future works

Conclusions
Tentative works

- ▶ Investigate the effects of other factors on nucleosome dynamics
- ▶ Develop an integrated framework for characterizing nucleosome dynamics from many kinds of data
- ▶ Elucidate the relationship between nucleosome dynamics, regulatory factors and gene regulation

Background
Characterizing nucleosome dynamics
Conclusions and future works

Conclusions
Tentative works

THANKS FOR YOUR ATTENTION!
Q&A
ngoctule@jaist.ac.jp