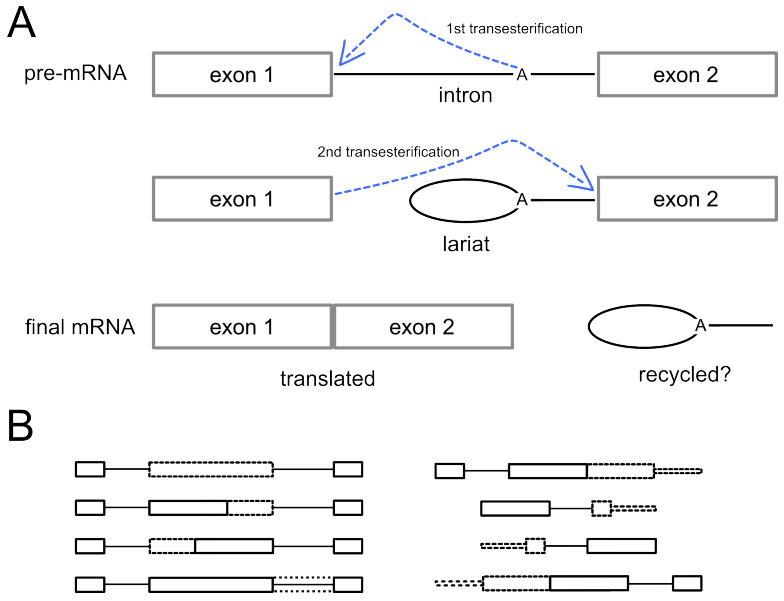
One exon, one nucleosome

Satu Nahkuri
Institute for Molecular Bioscience
The University of Queensland
2 Jun 2010

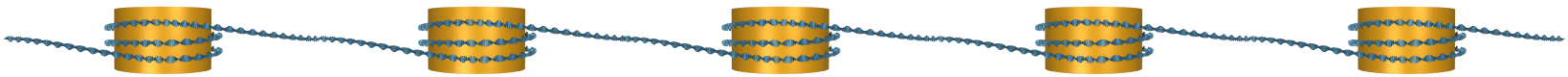
The curious case of alternative splicing

- Major contributor to proteome diversity
- Mechanism of tissue-specificity
- High fidelity despite numerous decoy splice sites
- The target sequences of auxiliary splicing factors are short, poorly conserved, or entirely lacking

- How is splicing regulated?

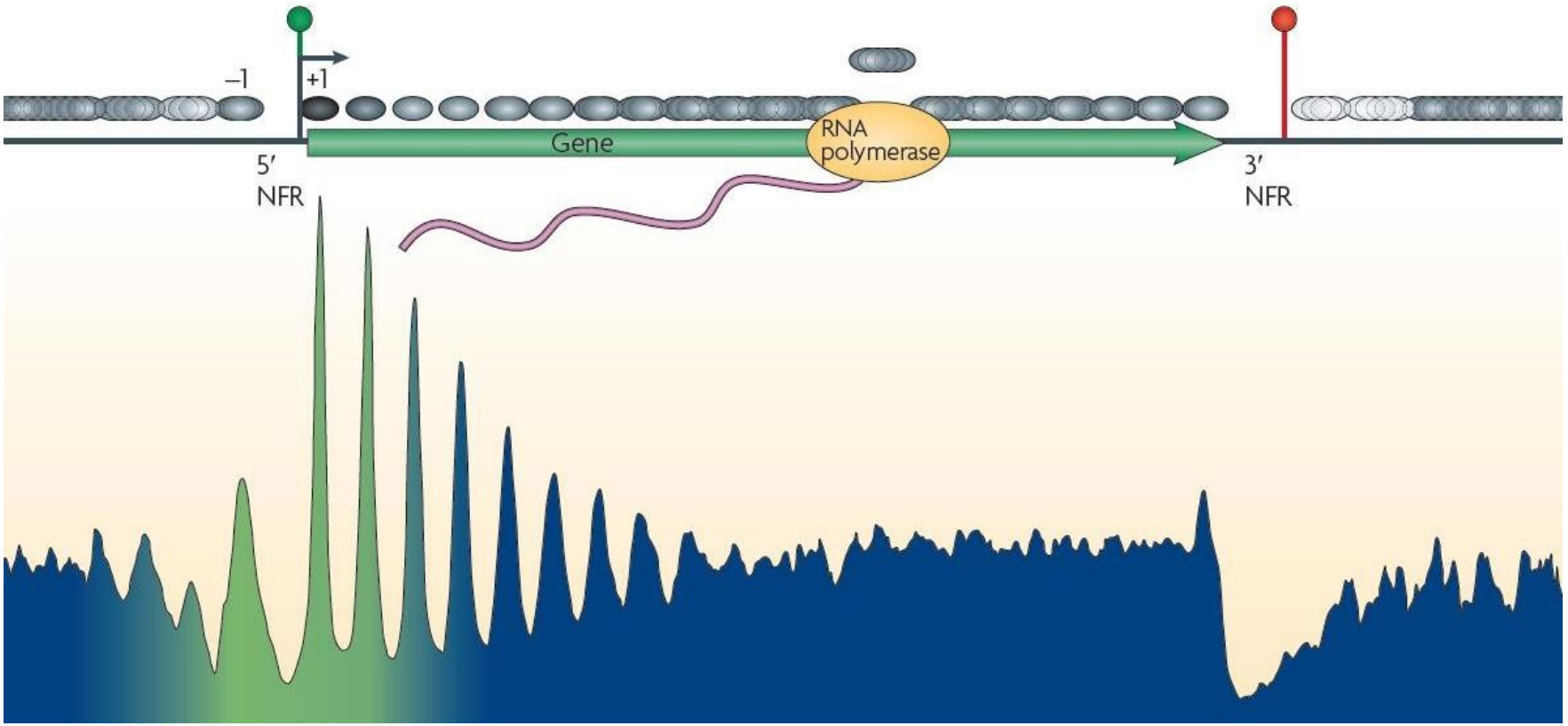


Nucleosomes in a nutshell



- One nucleosome consists of ~147 bp genomic DNA wrapped around a histone octamer
- Separated by ~10-80 bp linker DNAs
- Tails of histone proteins are subject to post-translational modifications that contribute to transcription initiation and elongation as well as alternative splicing
- Transcription, splicing and chromatin structure are functionally intertwined

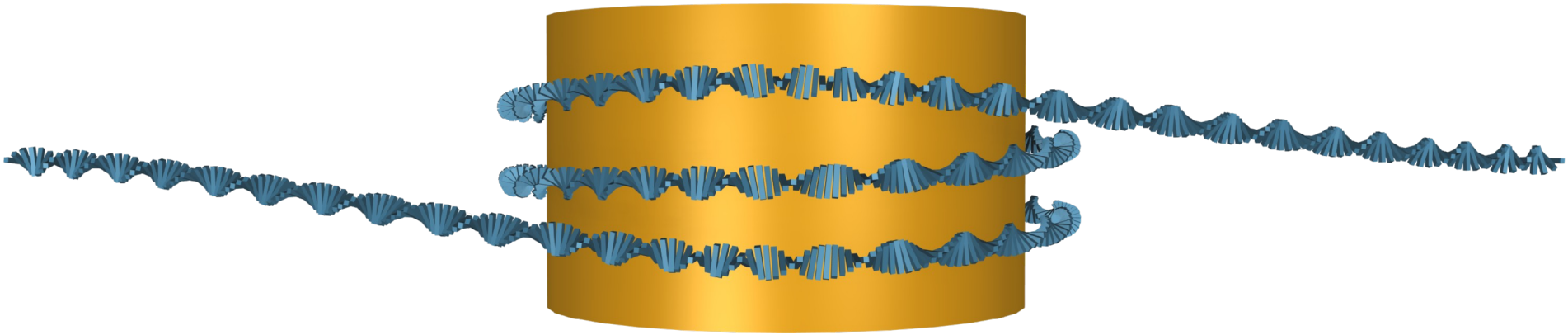
Nucleosome positioning and the gene



- Statistics from Jiang and Pugh, 2009 (data from yeast, but fly and human have the same features)
- Nucleosome free regions (5’ NFR, 3’ NFR)
- Nucleosome phasing ~1000 bp downstream TSS

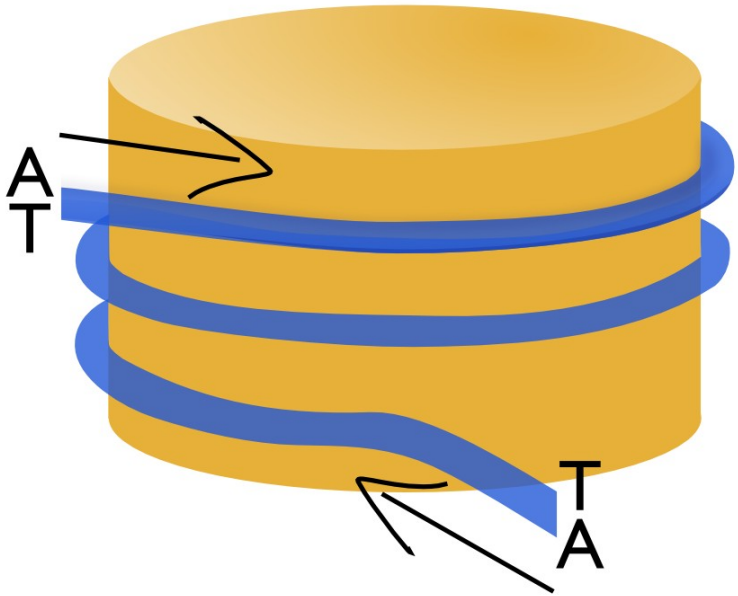
Nucleosomes at splice sites

- Primary sequences of exons and splice sites have a ~10.5 bp periodicity reminiscent of the pattern associated with fixed rotational phasing of nucleosomes (Baldi *et al*., 1996; Kogan & Trifonov, 2005)
- Two counterphase patterns:
  - Dinucleotides contracting the minor groove -> nucleosome surface
  - Dinucleotides expanding the minor groove -> outer rim of DNA



Genome-wide nucleosome positioning maps in CD4+ T cells

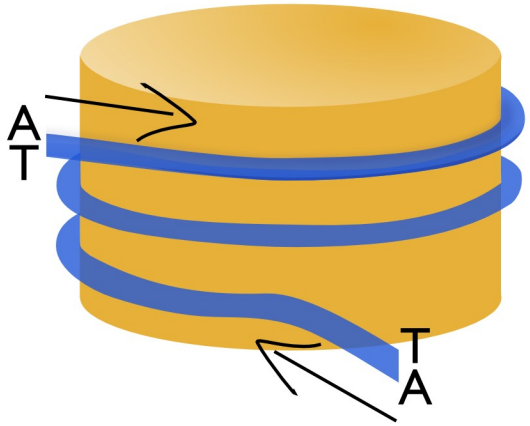
- Transcription start site studies by Barski, Schones *et al*.
- Linker DNAs are cut and consumed with MNase treatment
- Up to 36 bp from the 5’ ends of the chromatin associated DNA strands is deep sequenced by Solexa and aligned with the reference genome
- Barski: ChIP-Seq
- Schones: just Seq



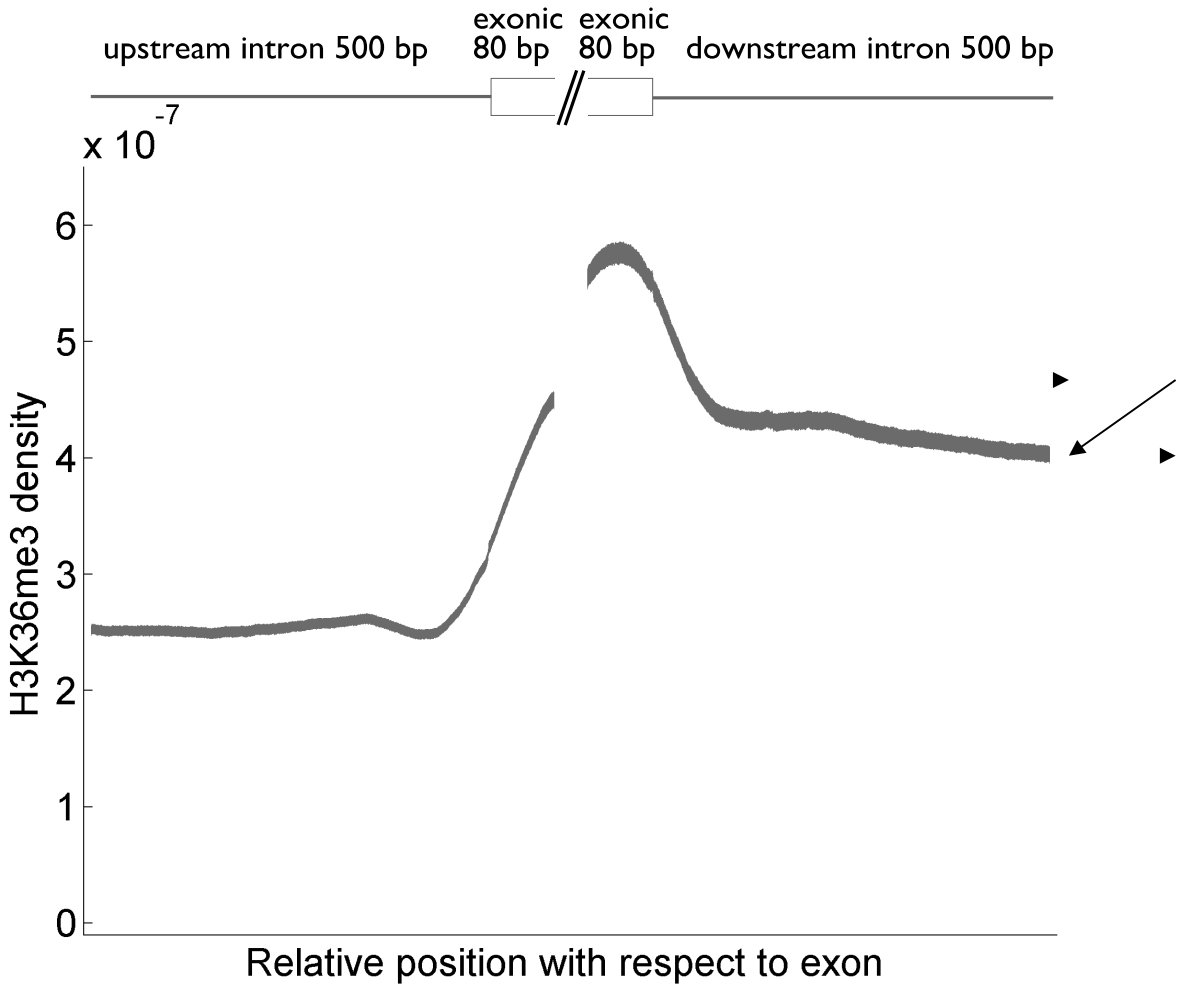
(Raw data from
Barski *et al.*, 2007,
Schones *et al.*, 2008)

Bioinformatic reassessment

- Extend the tags *in silico* to 150bp
- Calculate statistics at fixed 5’ and 3’ splice sites
- Analyze separately
  - Various histone modifications (ChIP-Seq)
  - Total nucleosome data (Seq without ChIP)



Positive control: H3K36me3



Original finding: Kolasinska-Zwierz *et al.*, 2009

- Excluded: 5' NFR+2kb, 3' NFR+1kb, all regions with any multimapping tags

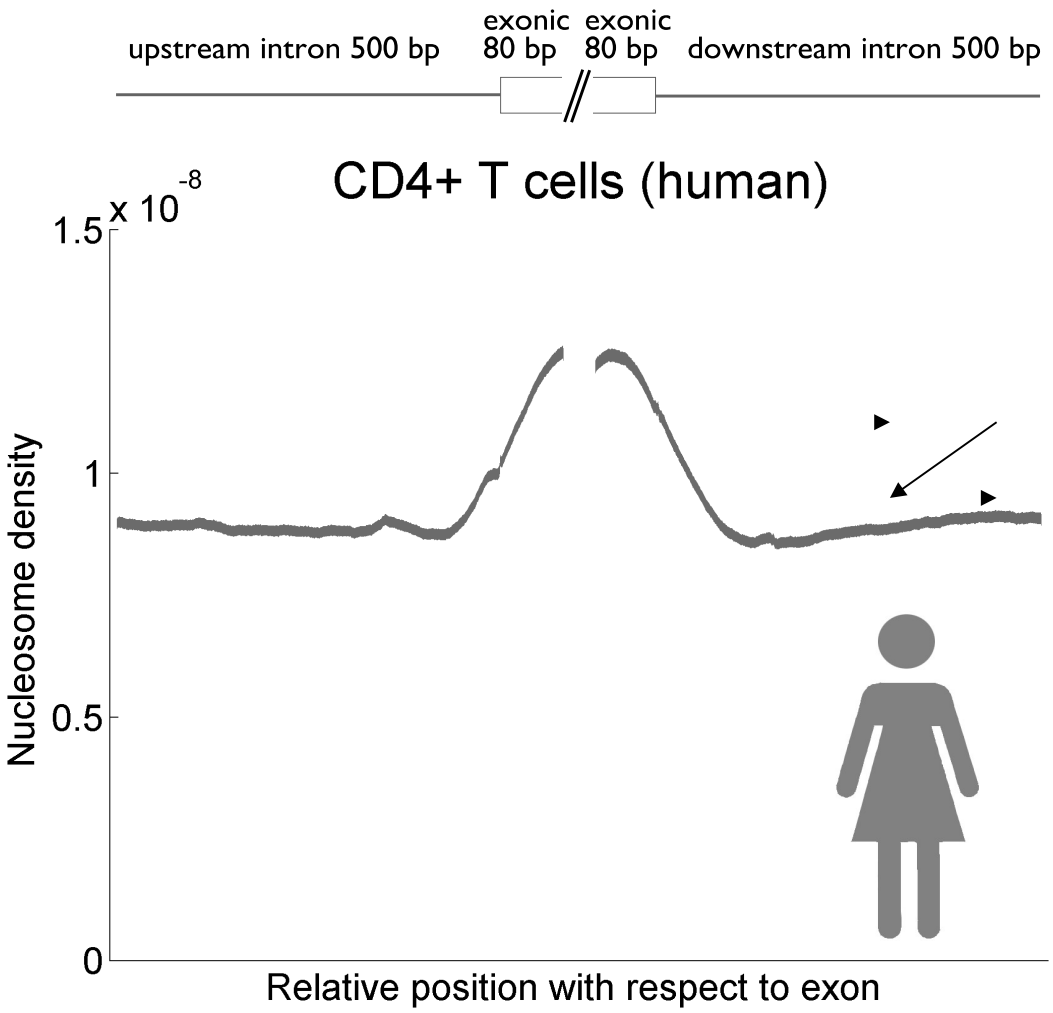
Research question

Density of H3K36me3 modified nucleosomes is higher at exons and downstream intronic regions of expressed genes – how about the density of other nucleosomes, modified or not?

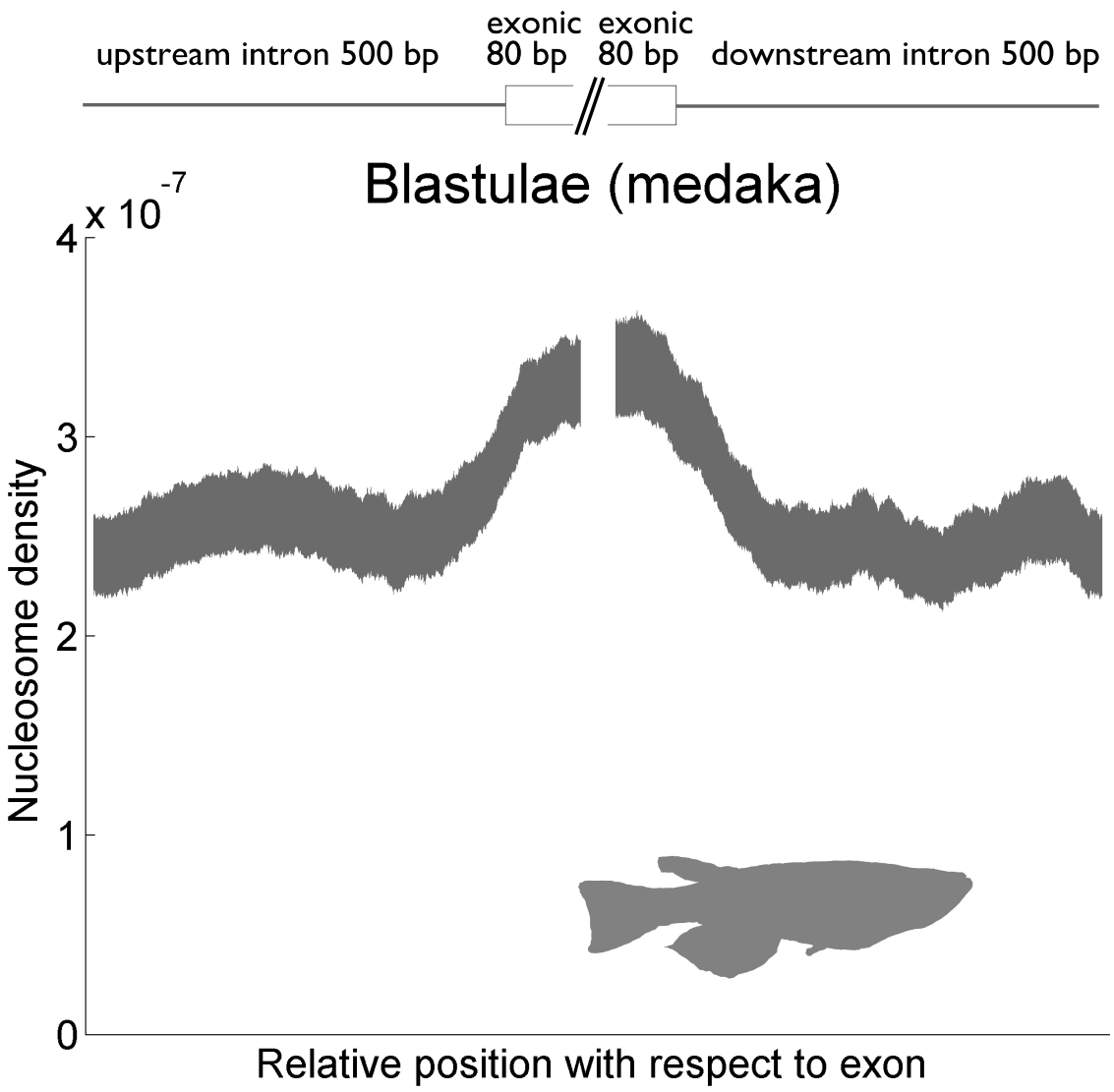
In contrast to earlier studies…

- Schones *et al.* were only investigating transcription start sites and polyadenylation sites and disregarded the rest of their nucleosome positioning library
- Kolasinska-Zwierz *et al*. did not study the total nucleosome density at exons at all

Nucleosomes are enriched at internal exons, but depleted at nearby introns

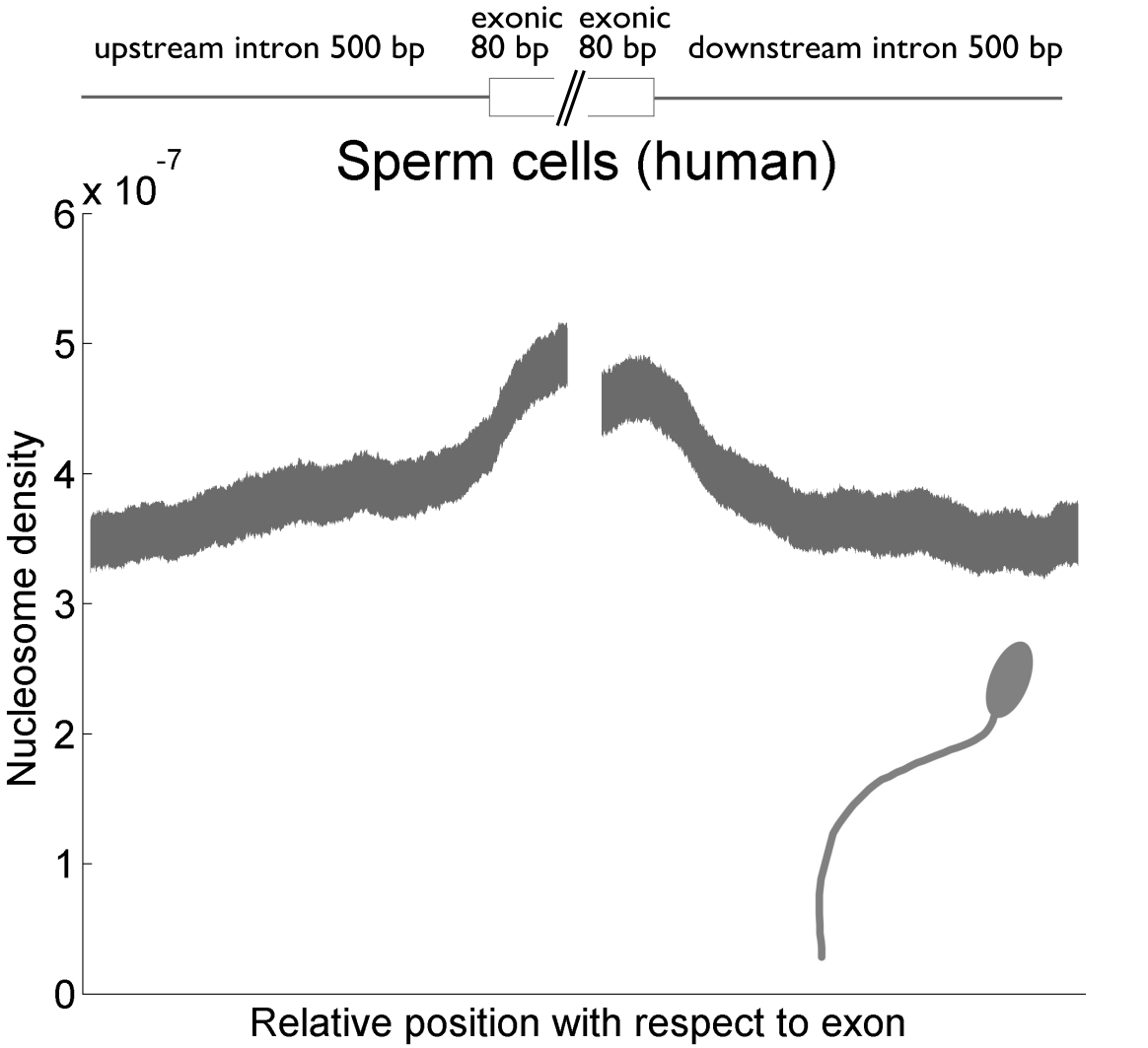


Enrichment is conserved in vertebrates



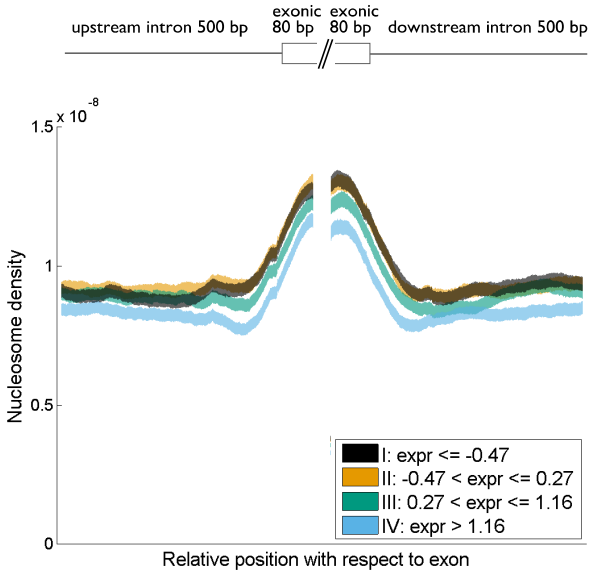
(Raw data from Sasaki *et al*., 2009)

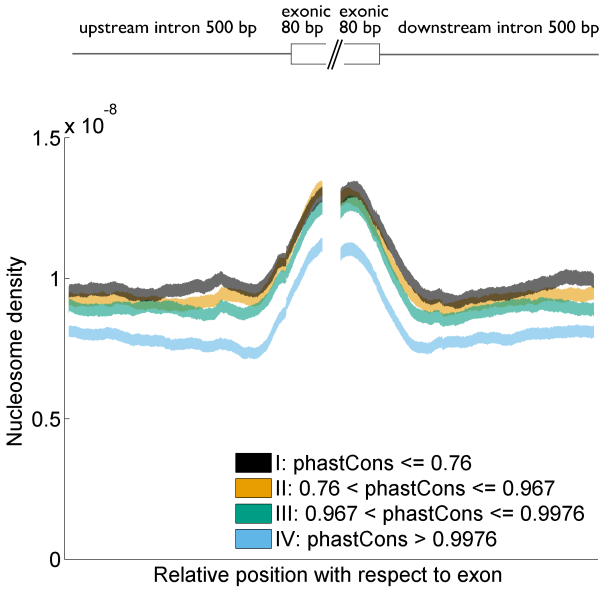
…and in sperm cells

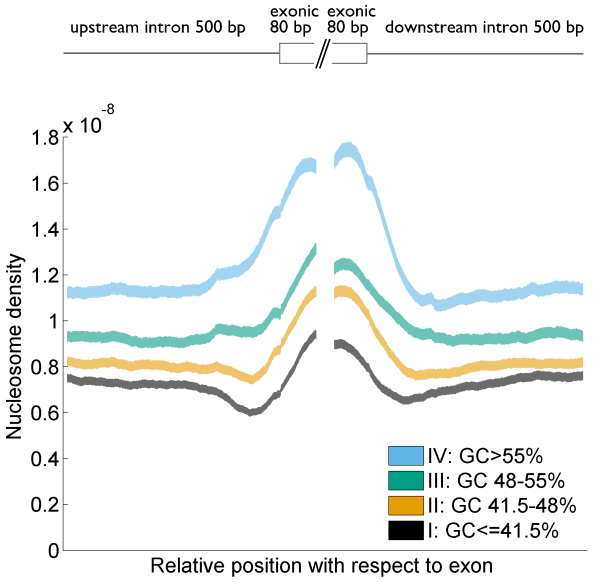


(Raw data from Hammoud *et al*., 2009)

Enrichment is independent of the expression level, conservation, and the GC content of the exon

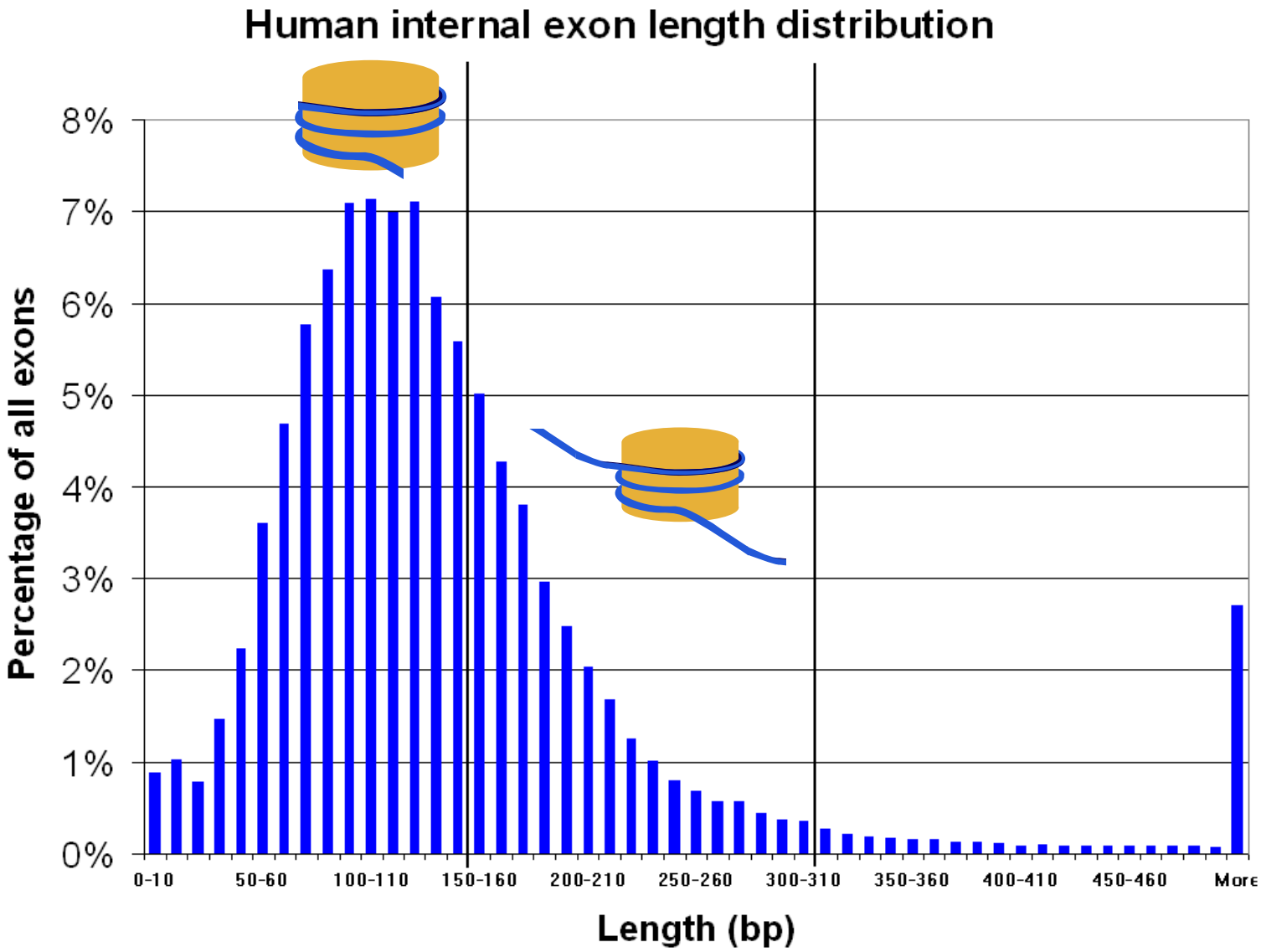






- Importantly, while the enrichment of H3K36me3 modified nucleosomes depends on the expression level, the total nucleosome enrichment can be observed *in the silent genes*

One exon, one nucleosome



Other studies

- Preferential positioning of nucleosomes at exons has been independently confirmed at least five times:
  - Andersson *et al*. 2009
  - Schwartz *et al*. 2009
  - Spies *et al*. 2009
  - Tilgner *et al*. 2009
  - Chen *et al*. 2010
- Both ChIP-seq and ChIP-chip data show the enrichment as well as the depletion at flanking intronic regions
- No signal at <50 bp exons, weak signal at 50-90 bp
- Nucleosome signal is stronger at weak splice sites
- No signal at pseudoexons

Conclusions

- Location of internal exons
  - is recorded in the chromatin structure
  - may be inherited across generations
- Epigenetic regulation operates at an exonic level, rather than only at a genic level
- A cornucopia of precious data awaits for bioinformatic reassessment in public genome-wide deep sequencing libraries

Thank you!

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

IMB Institute for Molecular Bioscience