Categorical spectral analysis of periodicity in nucleosomal DNA

Hu Jin[1, 2], H. Tomas Rube[1, 2, 3], Jun S. Song[1, 2, 4]

Center for the Physics of Living Cells

[1]Department of Physics, University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA, [2]Institute for Genomic Biology, University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA, [3]Present address: Department of Biological Sciences, Columbia University, New York, NY 10027, USA, and [4]Department of Bioengineering, University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA

Introduction

Nucleosome, consisting of 147 basepairs (bps) of DNA wrapping around a histone octamer, is the fundamental sub-unit of eukaryotic chromatin (Fig. 1A). The precise genomic location of nucleosomes plays a central role in local chromatin structure formation and epigenetic gene regulation. Certain 10.5-bp periodic nucleotides have been suggested to facilitate nucleosome positioning (Fig. 1B). However, key questions regarding the extent of nucleotide periodicity in nucleosomal DNA and its significance in directing nucleosome positioning still remain poorly understood:

- What is the level of 10.5-bp periodicity in individual nucleosomal sequences, rather than in average nucleotide frequency of dyad-aligned sequences? (Fig. 1C)
- How does periodicity affect single nucleosome positioning *in vivo*?

We address these questions by applying categorical spectral analysis to high-resolution nucleosome maps in yeast[1,2].

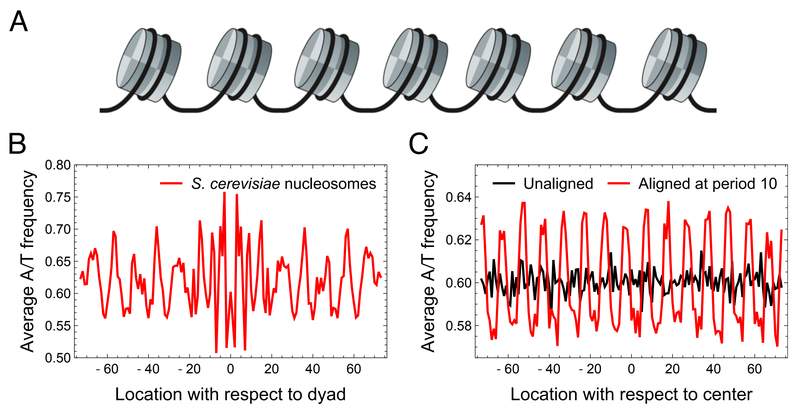

Fig. 1: Periodicity in dyad-aligned nucleosomal sequences does not imply periodicity in individual sequences.

(A) Illustration of nucleosomes, showing DNA wrapping around histone octamers (adapted from [3]). **(B)** A/T frequency averaged across dyad-aligned nucleosomal sequences. **(C)** A 10-bp periodicity can be created from 10,000 random sequences by systematically shifting the sequences by at most 5 bp.

Spectral envelope quantifies periodicity in individual nucleosomal sequences

Spectral envelope[4] explores all possible representations of DNA as real numbers simultaneously and computes the maximum spectral density among all possible representations. Applying spectral envelope to nucleosomal sequences in yeast showed that only a small amount of nucleosomal sequences contained excess 10.5-bp periodicity (15-20%, 4-5% significant at 5% FDR) (Fig. 2).

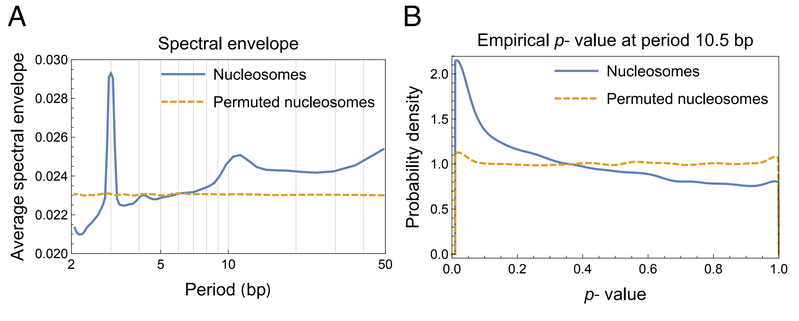

Fig. 2: Nucleosomal sequences in *S. cerevisiae* have enriched 10.5-bp periodicity compared to randomly permuted sequences.

(A) Average spectral envelope of nucleosomal sequences and randomly permuted sequences. **(B)** Distribution of p-values assessing the statistical significance of 10.5-bp periodicity in nucleosomal sequences and randomly permuted sequences.

Spectral decomposition quantifies origin of periodicity in dyad-aligned sequences

Both periodicity in individual sequences and proper phasing among sequences may contribute to the observed 10.5-bp periodic average nucleotide frequencies in dyad-aligned nucleosomal sequences. We developed a novel spectral decomposition method for quantifying the origin of observed periodicity in average nucleotide frequency (Fig. 3A). Applying the method to nucleosomal sequences in yeast demonstrates that the observed 10.5-bp periodicity in average nucleotide frequencies mainly stems from phasing among nucleosomal sequences (Fig. 3B, C).

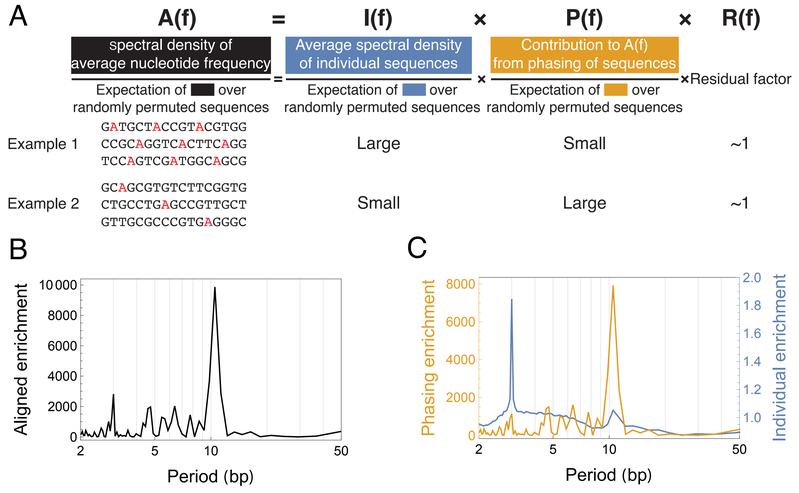

Fig. 3: Spectral decomposition of periodicity in average nucleotide frequency of dyad-aligned nucleosomal sequences.

(A) Illustration of the spectral decomposition. **(B)** The spectrum of aligned enrichment for nucleosomal sequences in *S. cerevisiae*, with A and T set to 1 and C and G to 0. **(C)** The spectrum of individual enrichment (blue) and phasing enrichment (yellow) in *S. cerevisiae*.

10.5-bp periodicity facilitates rotational but not translational positioning

Translational positioning of nucleosome refers to the location of the 147-bp DNA contacting the histone octamer; and rotational positioning refers to the rotational orientation of DNA helix relative to the histone surface[5]. To characterize the degree to which the translational location and the rotational orientation of a nucleosome vary across cells, we defined fuzziness scores for translational and rotational positioning, respectively. Integrative analysis of these fuzziness scores with spectral envelope demonstrates that 10.5-bp periodicity facilitates the rotational positioning of nucleosomes (Fig. 4A) but not translational positioning (Fig. 4B).

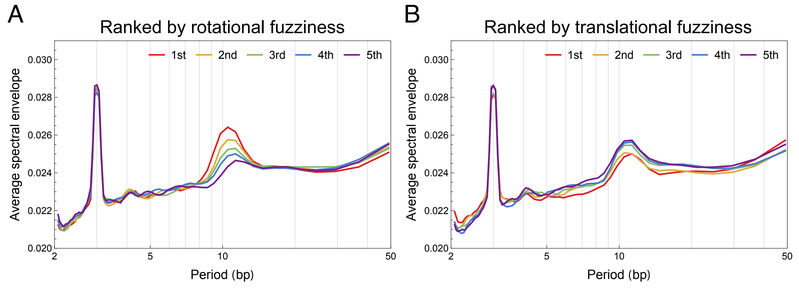

Fig. 4: Spectral envelope of nucleosomes grouped by the level of **(A)** rotational fuzziness and **(B)** translational fuzziness.

Nucleosomes were ranked from small to large values by rotational and translational fuzziness, respectively. The average spectral envelope of sequences within each quintile was then plotted, where the first quintile contains the smallest fuzziness.

Reduced 10.5-bp periodicity is a conserved signature of nucleosome-depleted regions around TSS/TTS

Rigid poly(dA:dT) tracts have been suggested as the main intrinsic signal in DNA sequences for creating nucleosome-depleted regions (NDR)[6] and are enriched in *S. cerevisiae* NDR (Fig. 5A). However, NDR upstream of transcription start site (TSS) in *S. pombe* is depleted of poly(dA:dT) tracts[7] (Fig. 5B), indicating that poly(dA:dT) tracts are not universally conserved features in NDR. Our spectral envelope analysis showed that the NDRs at TSS and transcription termination sites (TTS) in both *S. cerevisiae* and *S. pome* contained reduced 10.5-bp periodicity (Fig. 5C, D), indicating a potentially conserved function of sequence periodicity in creating NDR.

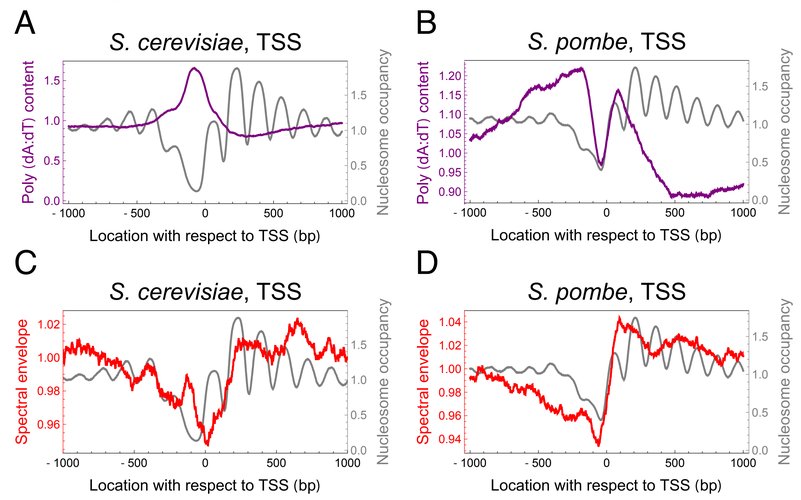

Fig. 5: Poly(dA:dT) contents and 10.5-bp periodicity around TSS in *S. cerevisiae* and *S. pombe*.

(A) Poly(dA:dT) content (normalized to the genome-wide mean) and nucleosome occupancy (normalized to the genome-wide mean) around TSS in *S. cerevisiae*. **(B)** Same as (A), but for *S. pombe*. **(C)** Spectral envelope at 10.5 bp (normalized to the genome-wide mean) and nucleosome occupancy (normalized to the genome-wide mean) aligned at the TSS in *S. cerevisiae*. **(D)** Same as (C), but for *S. pombe*.

Conclusions

Applying our rigorous computational framework based on categorical spectral analysis shows that

- Only a small fraction of individual nucleosomes in yeast actually contain a detectable 10.5-bp periodic pattern in nucleotide content.
- The previously observed periodicity in dyad-aligned nucleotide frequencies arises mostly from phasing of nucleosomal sequences.
- 10.5-bp periodicity, when present, significantly facilitates rotational, but not translational, nucleosome positioning.
- Reduced periodicity is an evolutionarily conserved signature of nucleosome-depleted regions around transcription start/termination sites.

The present framework can be easily adapted and applied to analyzing genome-wide nucleosome maps generated by various experimental techniques.

References

1. Brogaard K., *et al. Nature* **486** (2012): 496-501.
2. Moyle-Heyrman, G., *et al. PNAS* **110** (2013): 20158-20163.
3. Meyer C. A., *et al. Nat. Rev. Genet.* **15** (2014): 709-721.
4. Stoffer D. S., *et al. Biometrika* **80** (1993): 611-622.
5. Widom J. *Q. Rev. Biophys.* **34** (2001): 269:324.
6. Suter B., *et al. Nucleic Acids Res.* **28** (2000): 4083-4089.
7. Lantermann A. B., *et al. Nat. Struct. Mol. Biol.* **17** (2010): 251-257.