Genome-wide Comparison of DNA Melting Bubbles and Ensembl's Genomic Annotations

Geir Ivar Jerstad[1], Eivind Tøstesen[2*], Torbjørn Rognes[1] and Eivind Hovig[1,2]
[1]Department of Informatics, University of Oslo, Norway
[2]Department of Tumor Biology, Rikshospitalet-Radiumhospitalet Medical Center, Oslo, Norway

INTRODUCTION

We propose an extension of the existing genomic annotations (genes, promoters, etc.), by adding new features that describe also the physical behavior of the genomic DNA molecule.

DNA melting algorithms predict the opening of bubbles along the sequence upon heating to 50-100 ºC. We have equipped the Ensembl genome browser with extra views [1] that show the calculated melting properties as stitch profile diagrams (see box) [2,3] and melting maps [4]. This enables us to explore the bubble organization visually and correlate it with the other genomic annotations.

DNA bubbles are the physical basis of the transcription and replication processes. In a broader perspective, this work is part of an effort to annotate and make browsable the (epi)genomic aspects of the physics and the 3D structure of chromatin DNA.

What is a Stitch Profile diagram?

A *stitch* represents the location of either a bubble or a helical region on the sequence. In this example, three alternative conformations of a melting DNA correspond each to a row of stitches. All stitches are merged into a *stitch profile* diagram, thus annotating the alternatives.
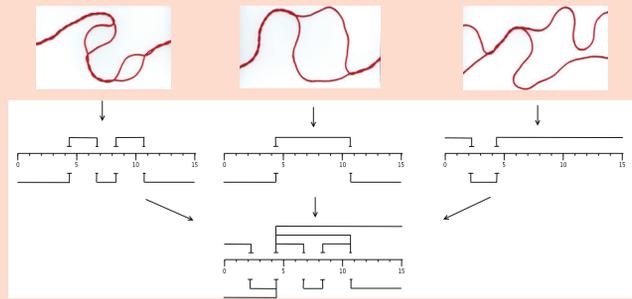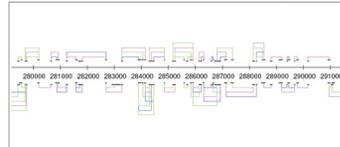

The complete stitch profile of yeast chromosome VIII (length 562,642 bps):
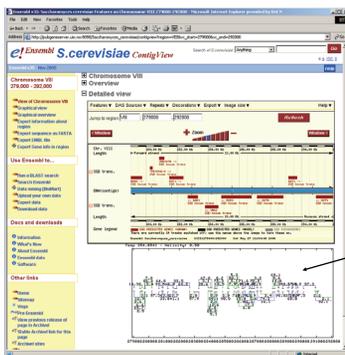

Screenshot from Ensembl including the stitch profile view, showing a partial alignment of yeast genes and stitches.

*Window 279k – 292k of the stitch profile.*

In this region, there are 13 predominant bubbles, shown as magenta-colored stitches (above the axis). Other colors indicate less probable conformations.

Screenshot from Ensembl including both a stitch profile view and a melting map view (i.e. Tm-profile).
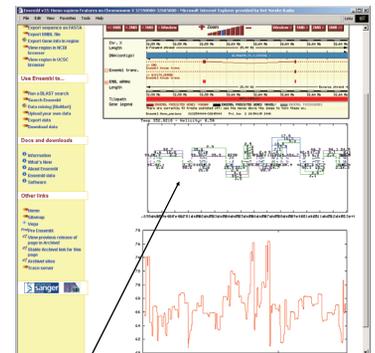
15k window: 32590000-32605000.

Stitch profile of the dystrophin gene (length 2.23 Mbps) on human chromosome X:
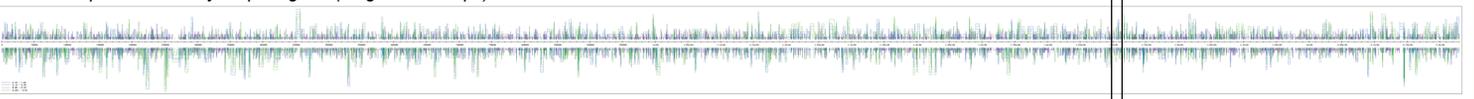
[1] G. I. Jerstad (2006): *Merging the physical properties of DNA with genomic annotations in Ensembl*. Master thesis, University of Oslo.
[2] E. Tøstesen (2005): Partly melted DNA conformations obtained with a probability peak finding method. *Phys. Rev. E*, 71, 061922.
[3] E. Tøstesen, G. I. Jerstad and E. Hovig (2005): Stitchprofiles.uio.no: Analysis of partly melted DNA conformations using stitch profiles. *Nucl. Acids Res*., 33, w573-w576. (Article describing our web server at http://stitchprofiles.uio.no.)
[4] F. Liu, E. Tøstesen, *et al* (2006): The human genomic melting map. *Manuscript in preparation*.

Comprehensive Cancer Center