

# Motif RNA maps: A short introduction

André Gohr

September 17, 2017

# Diff. AS caused by diff. binding of RBPs?

## Context

- ▶ given RBP: Nova
- ▶ experiment: Nova KD vs. WT
- ▶ after PSI estimation:  
**diff. spliced SEs vs non-diff. spliced SEs**

## Hypothesis

- ▶ diff. binding of RBP in exons or proximity causes diff. splicing

## Questions

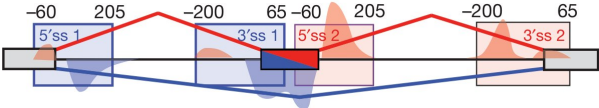
- ▶ enrichment of RBP binding sites?
- ▶ differences in location of binding sites?

# RNA maps: A visualization for studying RBP binding

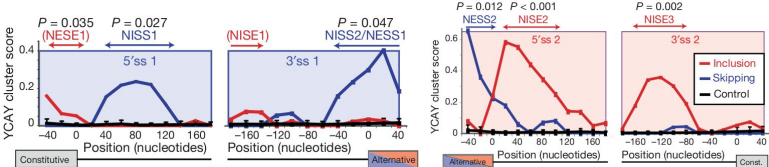
## RNA map ...

- ▶ ... visualization technique for visually inspecting enrichment and location bias of RBP binding in proximity of SE
- ▶ ... introduced by Ule et al. 2006: An RNA map predicting Nova-dependent splicing regulation

## Regions of interest (Ule et al.)



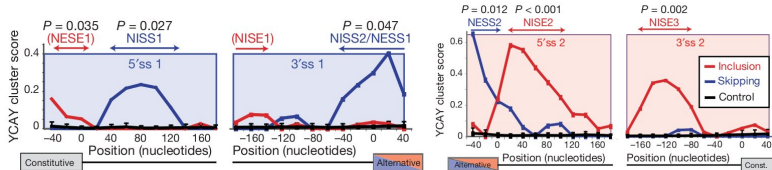
## Example (Ule et al.)



# Nomenclature

- ▶ not yet established widely, different terms used
- ▶ **RNA map**: mapping of some feature across proximity of SEs
- ▶ depending on feature, e.g.
  1. motif RNA map
  2. R loops RNA map
  3. bendability RNA map
  4. ...

## Motif RNA map



## Producing motif RNA maps with ...

- ▶ **rMAPS** website: <http://rmaps.cecsresearch.org>
  1. species: human, mouse, fly, worm, fish
  2. motifs: (not sure if any) Perl regular expressions (REGEXP)
  3. tied to rMATs
  4. exon centric (SE)
  
- ▶ **RBPmap** website: <http://rbpmap.technion.ac.il>
  1. species: human, mouse, fly
  2. motifs: IUPAC consensus (limited REGEXPs), RBPmap motifs
  3. exon centric (SE)
  
- ▶ **Matt**: stand-alone program
  1. species: any with available genome assembly
  2. motifs: any PWM, any Perl regular expression
  3. exon- and intron-centric (SE, IR)

## RNA maps with Matt: command rna\_maps

**Coordinates:** C1, START, END, C2

exon centric



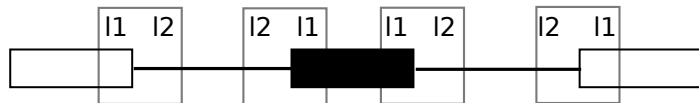
intron centric



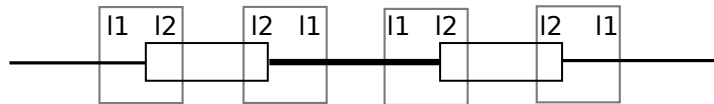
## RNA maps with Matt: command rna\_maps

**Exon / intron parts to be scanned: l1, l2**

exon centric



intron centric



## RNA maps with Matt: command `rna_maps`

**Length of sliding window:  $lw$**

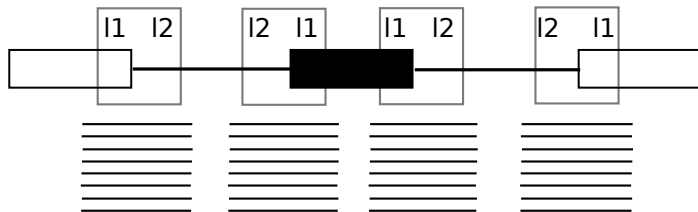
exon centric





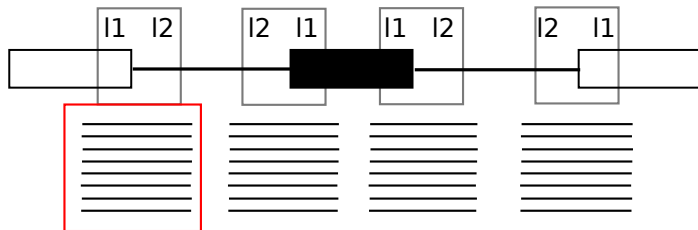
## Matt rna\_maps: 1.) sequence extraction

exon centric



## Matt rna\_maps: 1.) sequence extraction

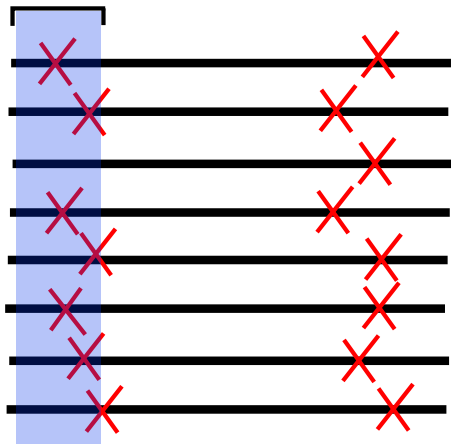
exon centric



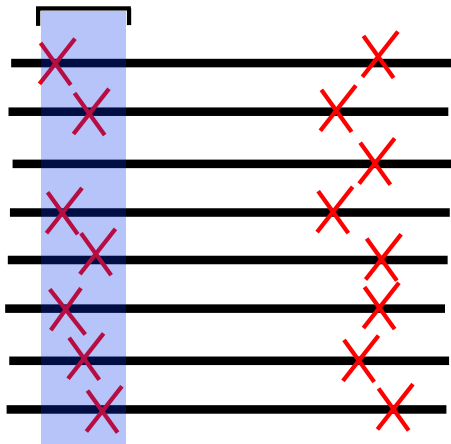
## Matt rna\_maps: 2.) search motif hits (PWM, REGEXP)



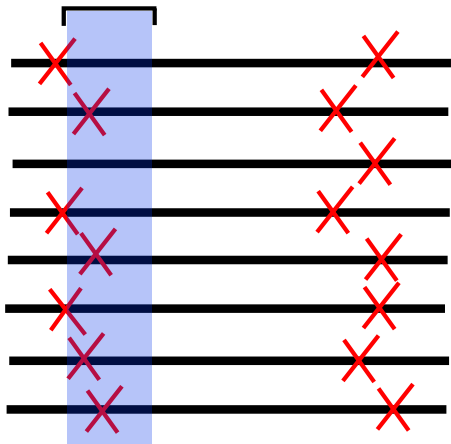
## Matt rna\_maps: 3.) move sliding window across sequences



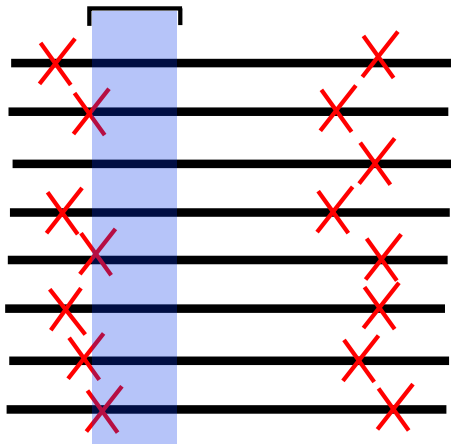
Matt rna\_maps: 3.) move sliding window across sequences



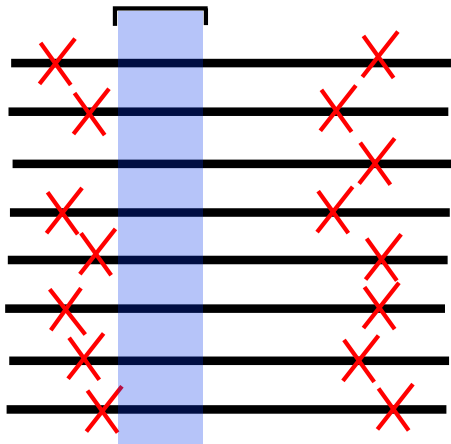
Matt rna\_maps: 3.) move sliding window across sequences



## Matt rna\_maps: 3.) move sliding window across sequences

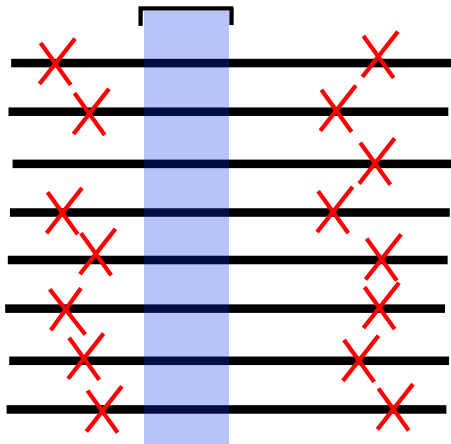


Matt rna\_maps: 3.) move sliding window across sequences

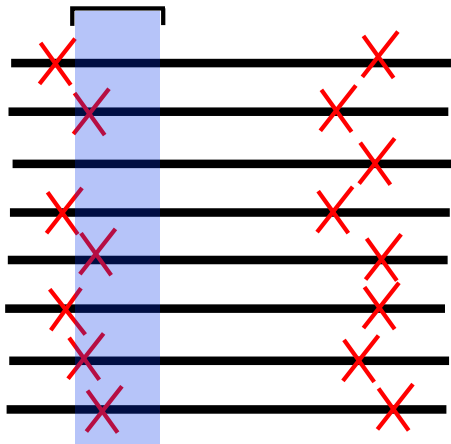




Matt rna\_maps: 3.) move sliding window across sequences



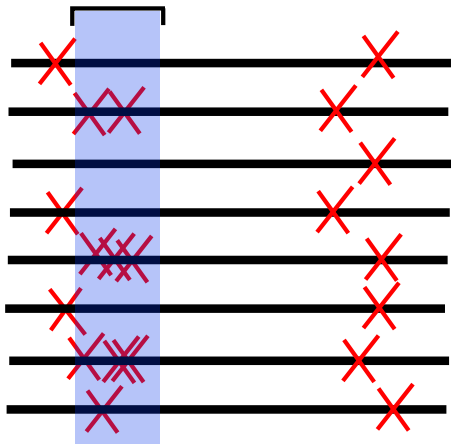
Matt rna\_maps: 4.) enrichment score for each win. position



Score 1: **Avg. number of hits per sequence:**

$$N \text{ hits} / N \text{ seqs} = 4/8$$

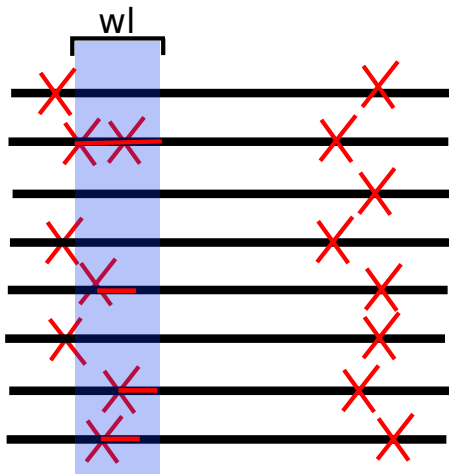
Matt rna\_maps: 4.) enrichment score for each win. position



Score 2: **Proportion of sequences with at least  $X$  hits:**

$$N \text{ seqs with 2 or more hits} / N \text{ seqs} = 3/8$$

Matt rna\_maps: 4.) enrichment score for each win. position



Score 3: **Fraction of region covered by motif:**

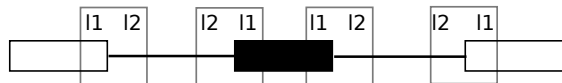
$$N \text{ covered pos.} / N \text{ seqs} \cdot wl = 5/16$$

if motif length is constant and  $wl/2$

## Matt rna\_maps: comparing groups of exons

Compare groups: determine enrichment score for each group and visualize in same RNA map

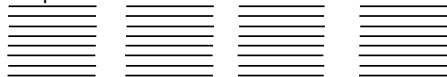
exon centric



diff. spliced SEs



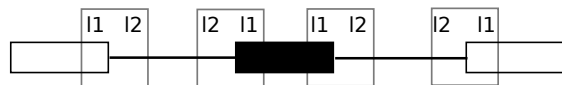
non-diff. spliced



## Matt rna\_maps: Nova2 motif RNA map

- ▶ mouse cortex: Nova2 KD vs. WT (3 replicates, vast-tools)
- ▶ 152 silenced ( $dPSI \geq 25$ ), 167 enhanced ( $dPSI \leq -25$ ), 322 unregulated exons

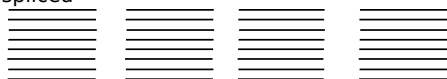
exon centric



up spliced SEs



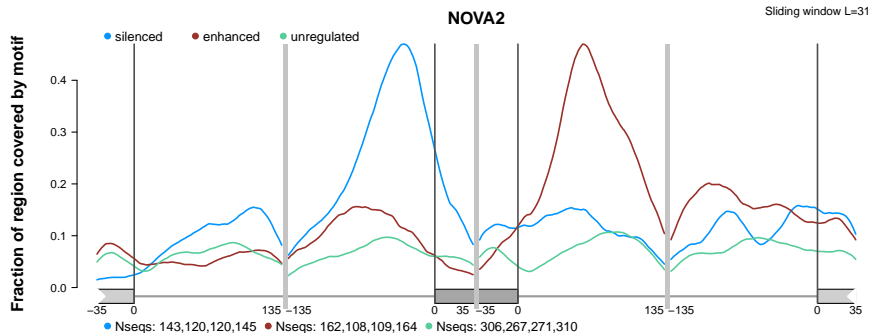
down spliced



non-diff spliced



# RNA rna\_maps: Nova2 motif RNA map



152 silenced, 167 enhanced, 322 unregulated exons

