

# Analysing RNA-seq data using GeneTiles

Jan de Sonnevile<sup>1#\*</sup>, Kees-Jan van der Kolk<sup>1,2\*</sup>, Wouter Veneman<sup>3</sup>, Robert Lodder<sup>2</sup>, Zaid Al-Ars<sup>2#</sup>, Herman Spaink<sup>3#</sup>

1) Life Science Methods BV, Leiden, the Netherlands

2) Computer Engineering, Delft University of Technology, Delft, the Netherlands

3) Molecular Cell Biology, Leiden University, Leiden, the Netherlands

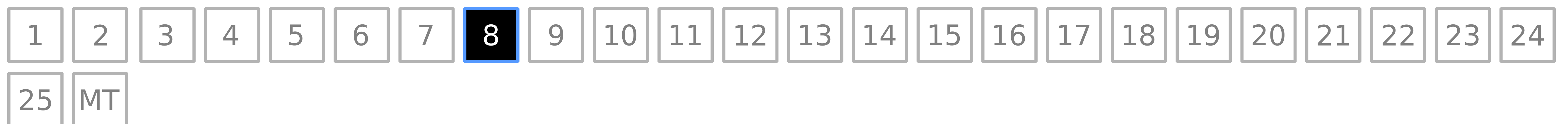
\*contributed equally to this work; #correspondence: jan@lifesciencemethods.com, Z.AI-Ars@tudelft.nl, h.p.spaink@biology.leidenuniv.nl

RNA-seq data, containing tens of millions of reads, is mostly processed using scripts. After processing, a selection of reads is analyzed using RNA-seq viewers. Directly browsing processed RNA-seq data is difficult due to the large dynamic range of length scales of reads (50bp), exons (~200bp), introns (~3kb), genes (~20kbp), and chromosomes (~65Mbp). In addition, current RNA-seq viewers show introns at the same scale as exons, which in most experiments means that 90 percent of the visible sequence data does not display aligned reads.

Using the data processing pipeline from a recent paper from our group, we created an online viewer which does allow for browsing all the aligned reads, while eliminating almost completely the need for user intervention (such as zooming in).

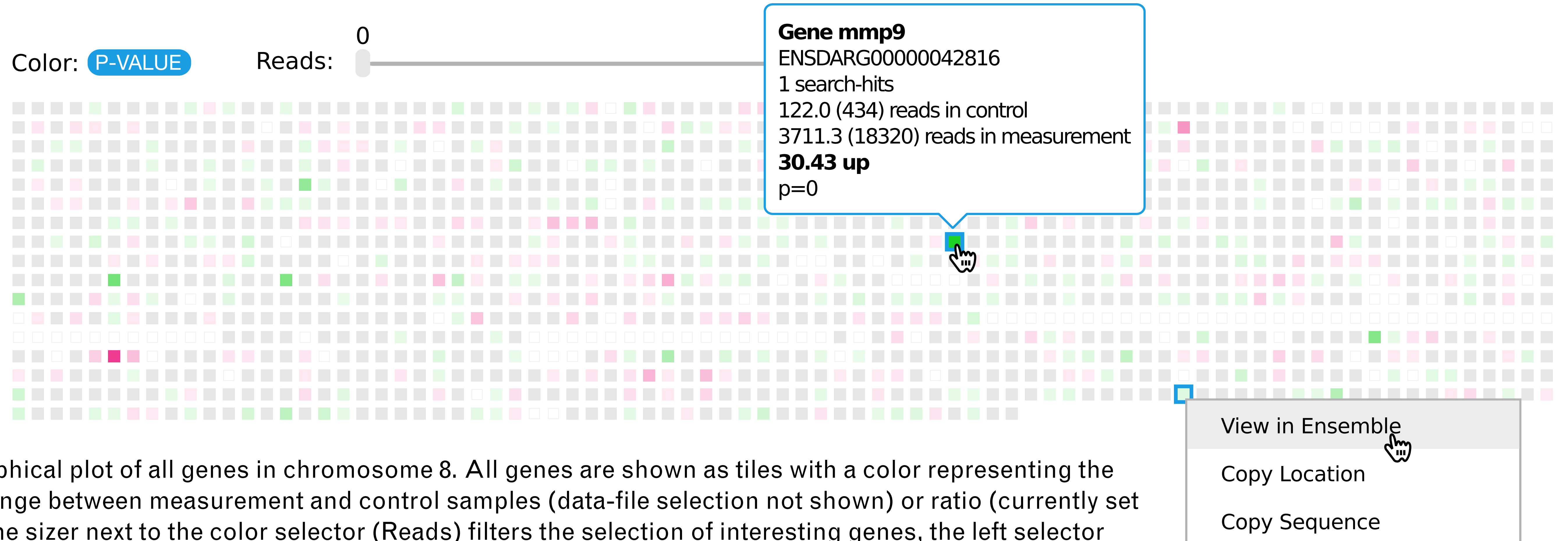
## Chromosomes

download ▾



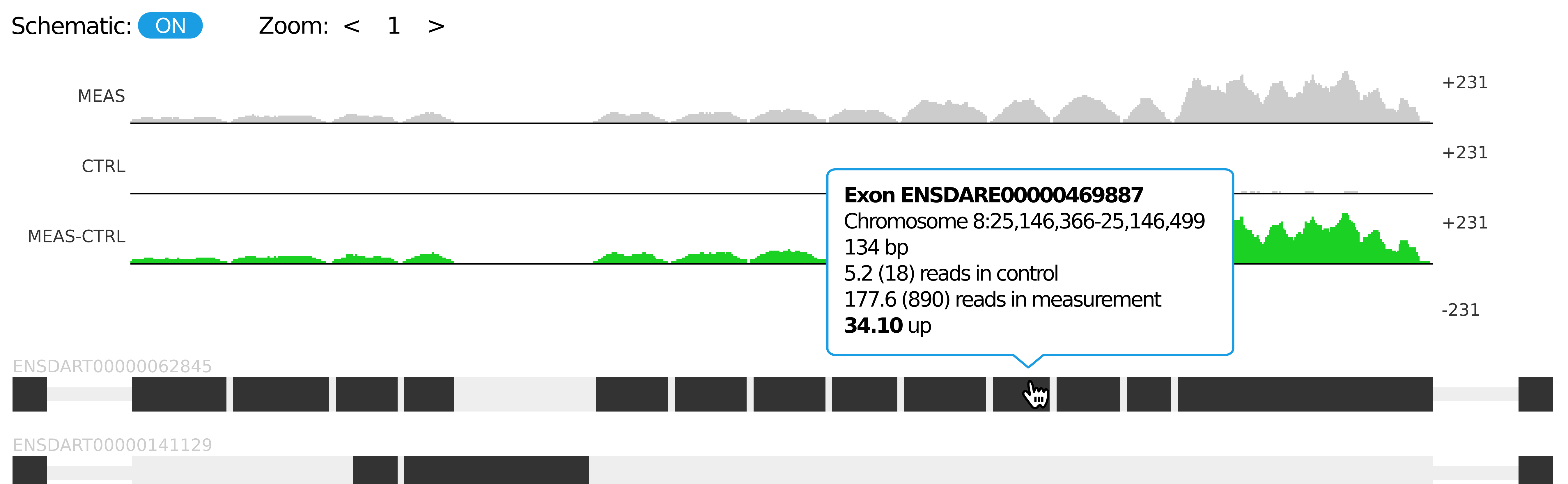
## Genes of chromosome 8

download ▾



**Figure 1:** Graphical plot of all genes in chromosome 8. All genes are shown as tiles with a color representing the p-value of change between measurement and control samples (data-file selection not shown) or ratio (currently set to p-value). The sizer next to the color selector (Reads) filters the selection of interesting genes, the left selector excludes genes with a minimum number of reads (currently 0) and the right selector (behind info-box) excludes a maximum number of reads (e.g. household genes). Here, green tiles represent significance of upregulated genes and red of downregulated genes. Unchanged expression of genes is shown in grey and white tiles represent genes either excluded from the selection or without reads. Mouse-movement over genes shows detailed information, and right-click allows to view this gene in Ensemble or export. Download (right side) offers the possibility to download all data in excel (ratios and p-values) or svg format (visualization as shown).

## Exons and Introns of gene *mmp9* / ENSDARG00000042816



**Figure 2:** From the selected gene an overview of the reads on exons and introns is shown as well as the difference, the different known transcripts are visible underneath. The schematic view shows the reads on the exons, by reducing introns to a minimal width. When schematic mode is off (not shown) the length between exons and introns is evenly scaled. In both views a minimal width per intron and exon is maintained.

GeneTiles will be available starting september 1, via <http://www.genetiles.com> for a live demo, info and early access, see booth of Life Science Methods BV