

What is all this genome expression?

Observations and statistics for expression at the
base level

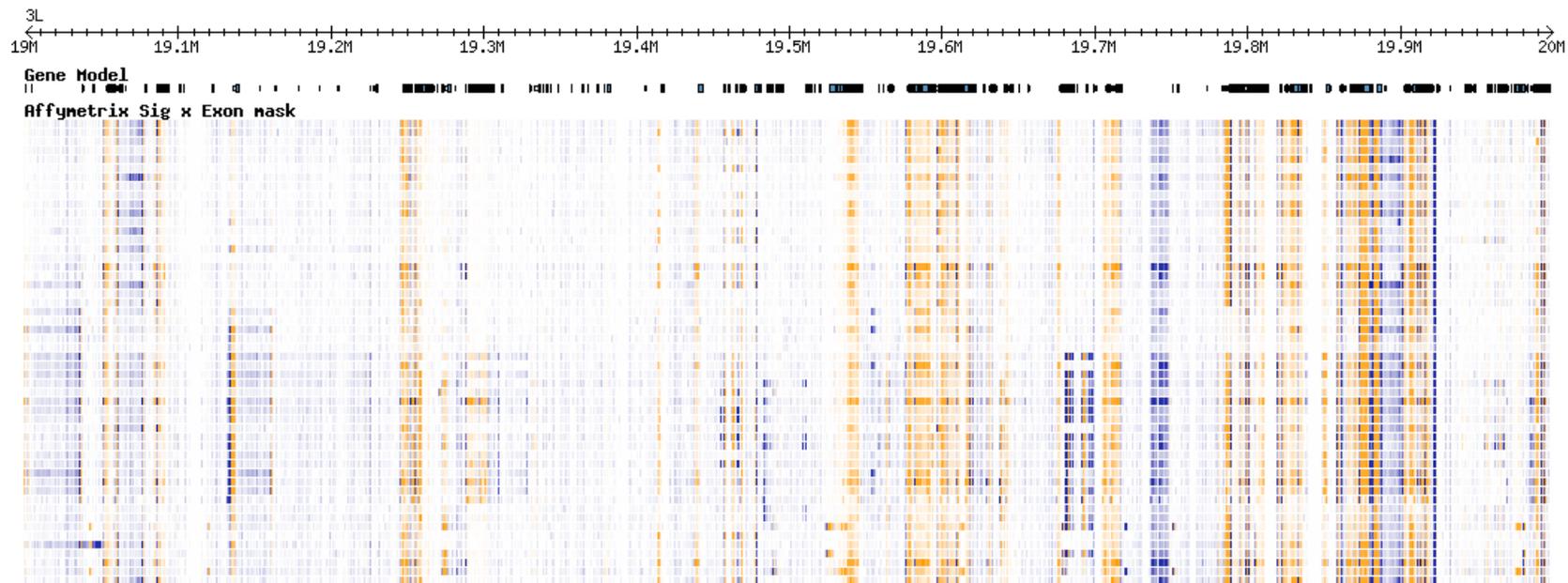
Don Gilbert

April 2009

Biology Dept., Indiana University
gilbertd@indiana.edu

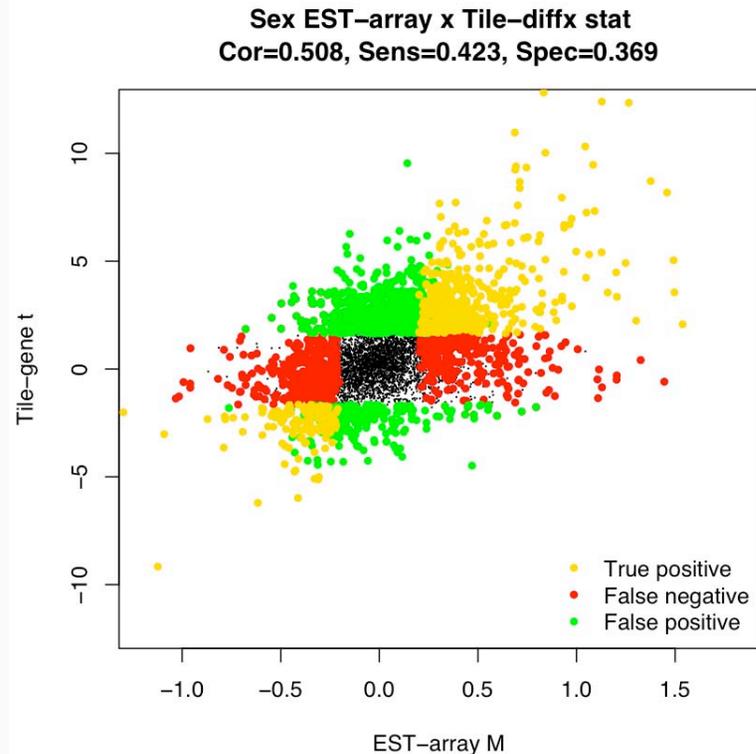
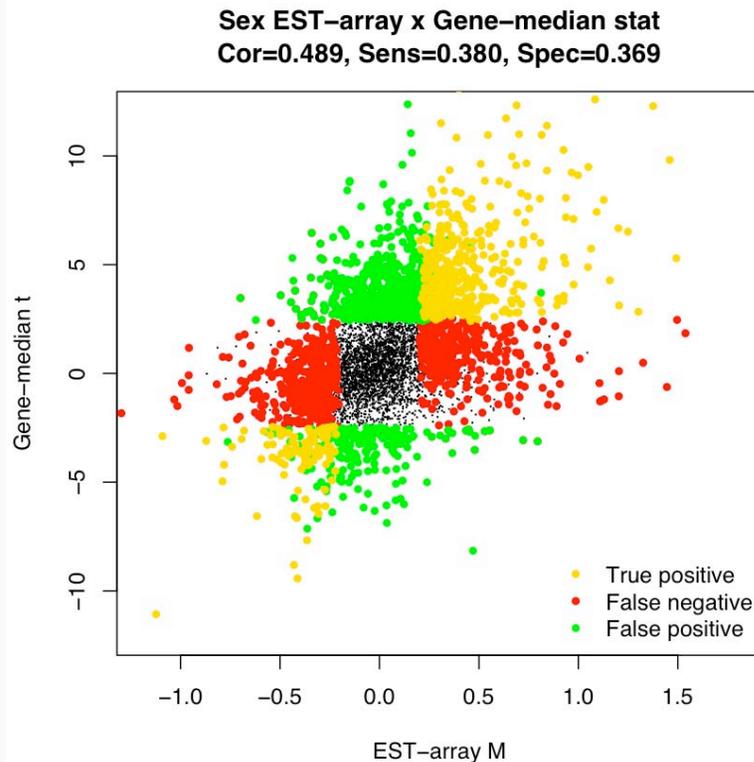
2007: Tile expression

DrosMel tiled by Affymetrix, finds new genes (blue) and known (orange)



Gene or Tiled expression?

Tiled genes find expression better than gene average

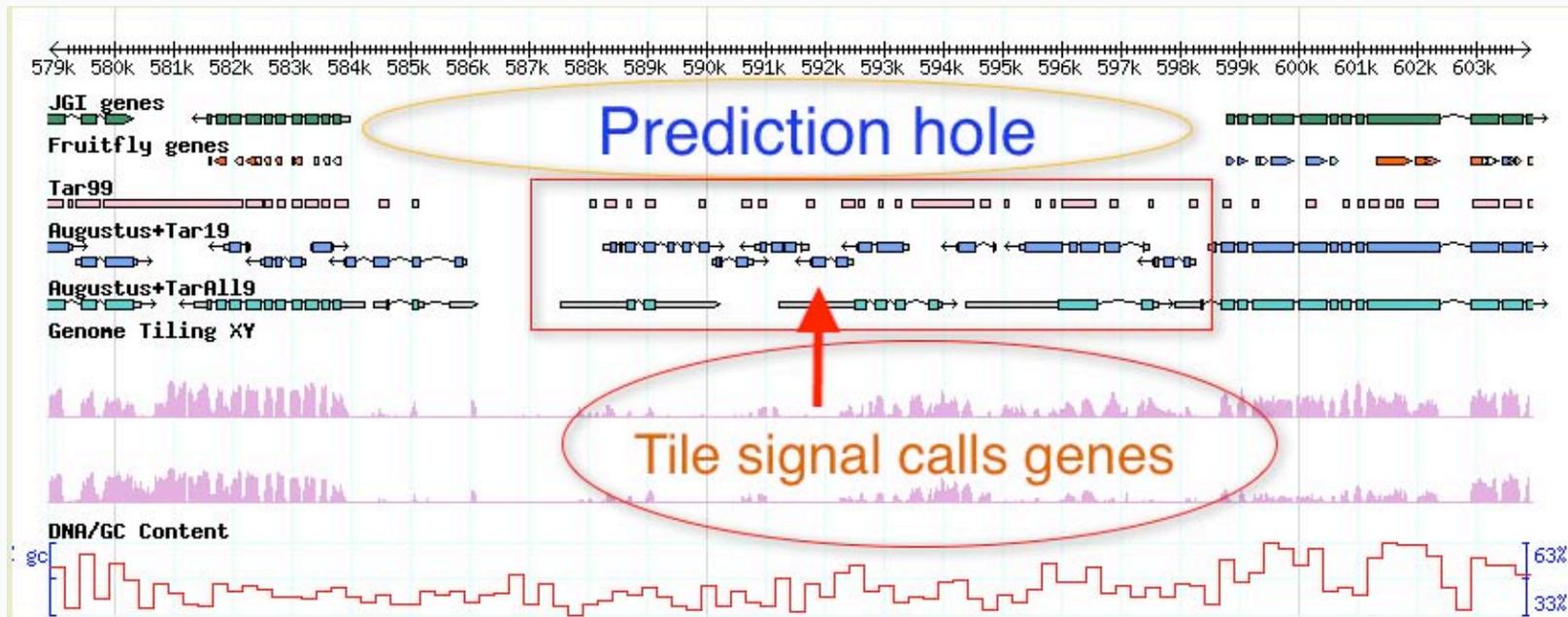


Gene median sensitivity 38%, specificity= 37%

Tiled gene sensitivity= 42%, specificity= 37%
in reference to separate sex expression study

Tiles find new genes

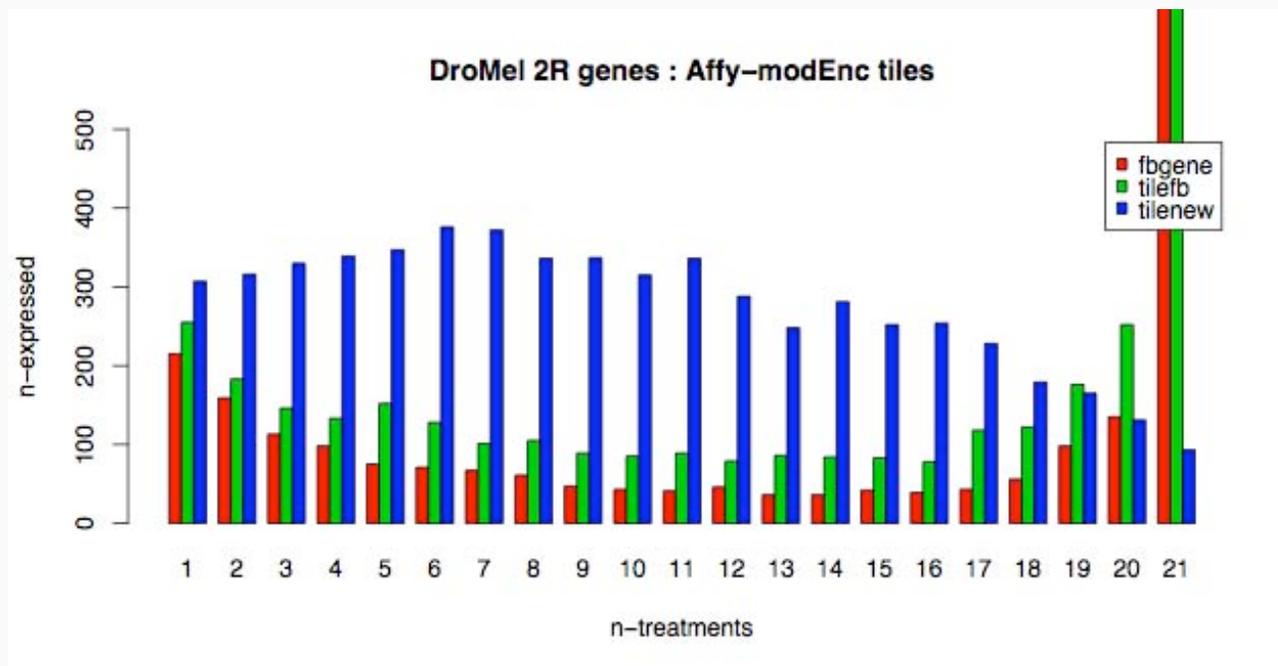
Daphnia tile expression with gene finding calls 26% coding bases over the genome, compared to 17% from gene predictions, or 5,000 - 10,000 new genes.



Manak *et al* 2006, with Drosnel also found 24% CDS/genome, up from 18% CDS/genome from reference gene set. Computational tools need to mature; gene finding is preliminary.

Tiles find rarely used genes

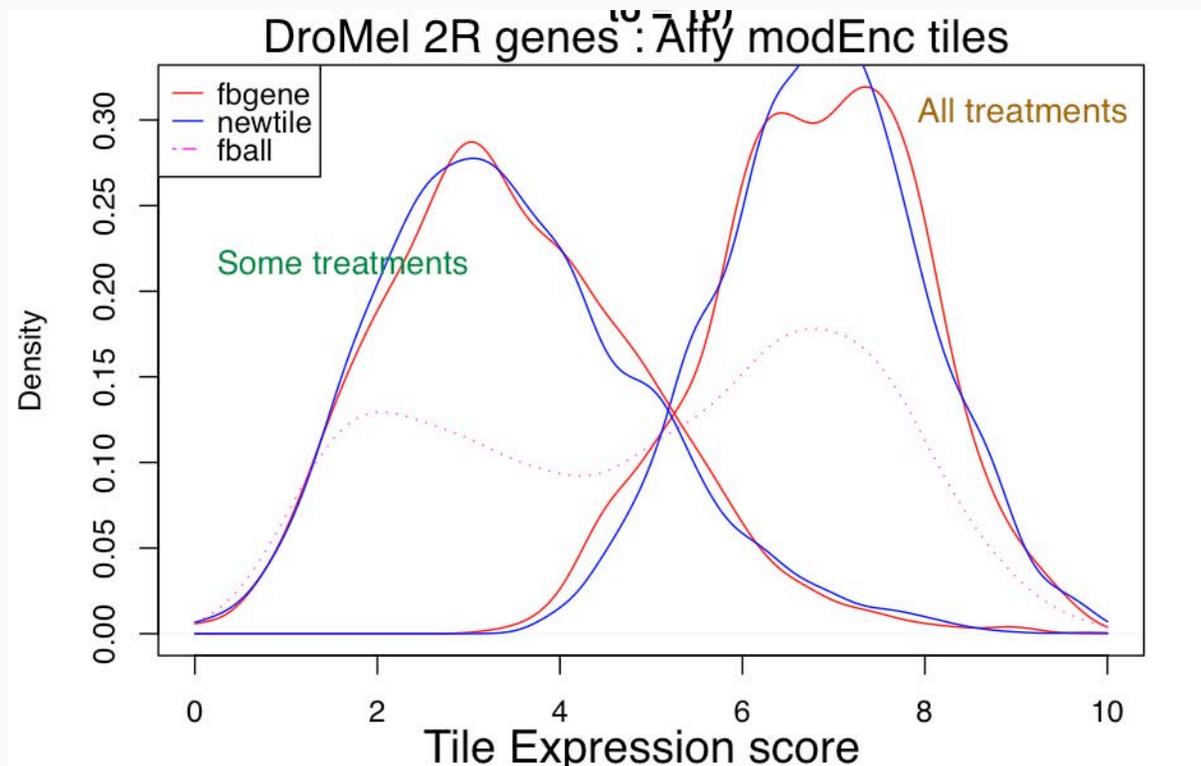
New tile-genes are expressed in a subset of conditions, most known genes are expressed in all conditions



But some known genes have same low, spotty expression as Tile-genes

Tiles find rarely used genes

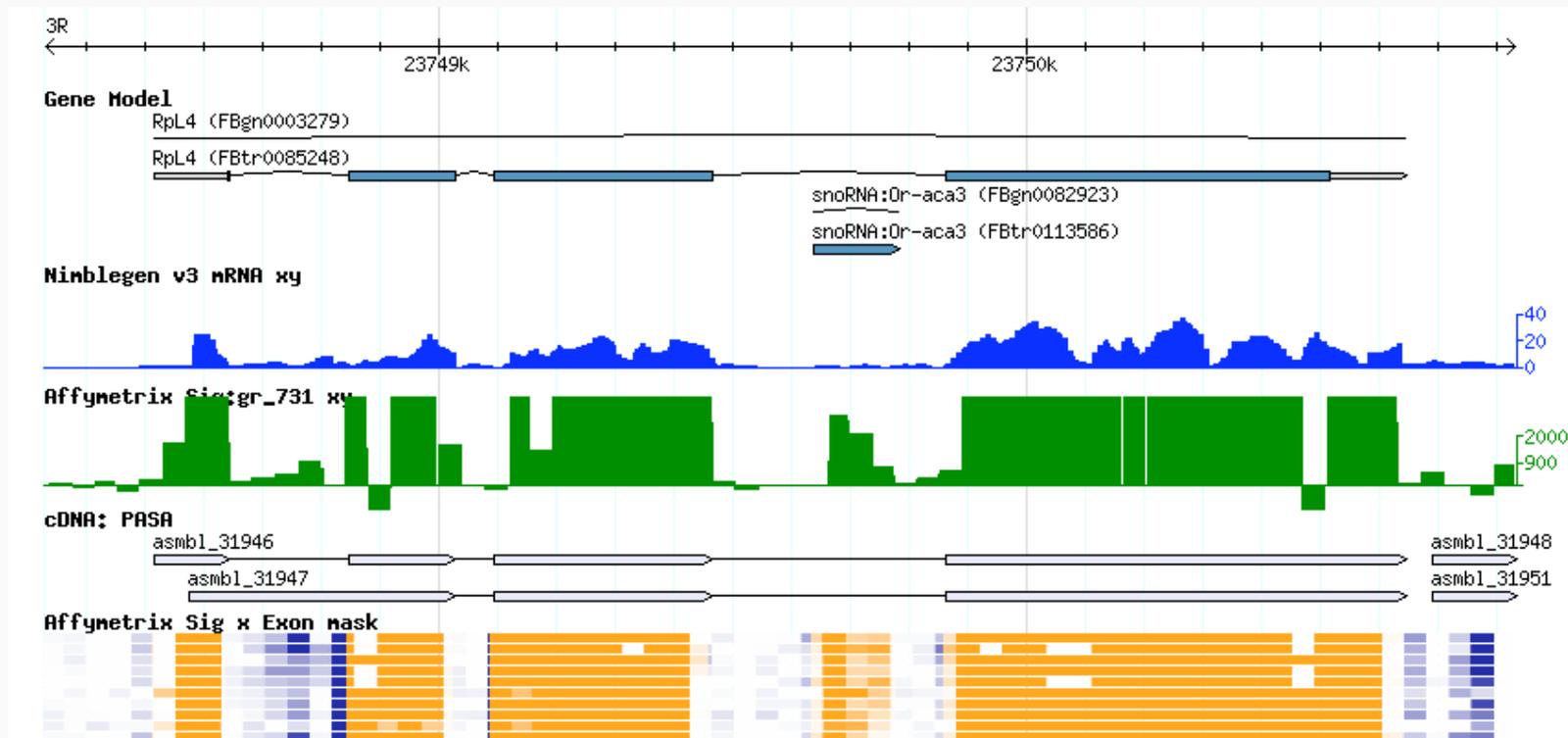
Average expression is high for genes found in all treatments



Accurate gene detection at low level expression remains a challenge

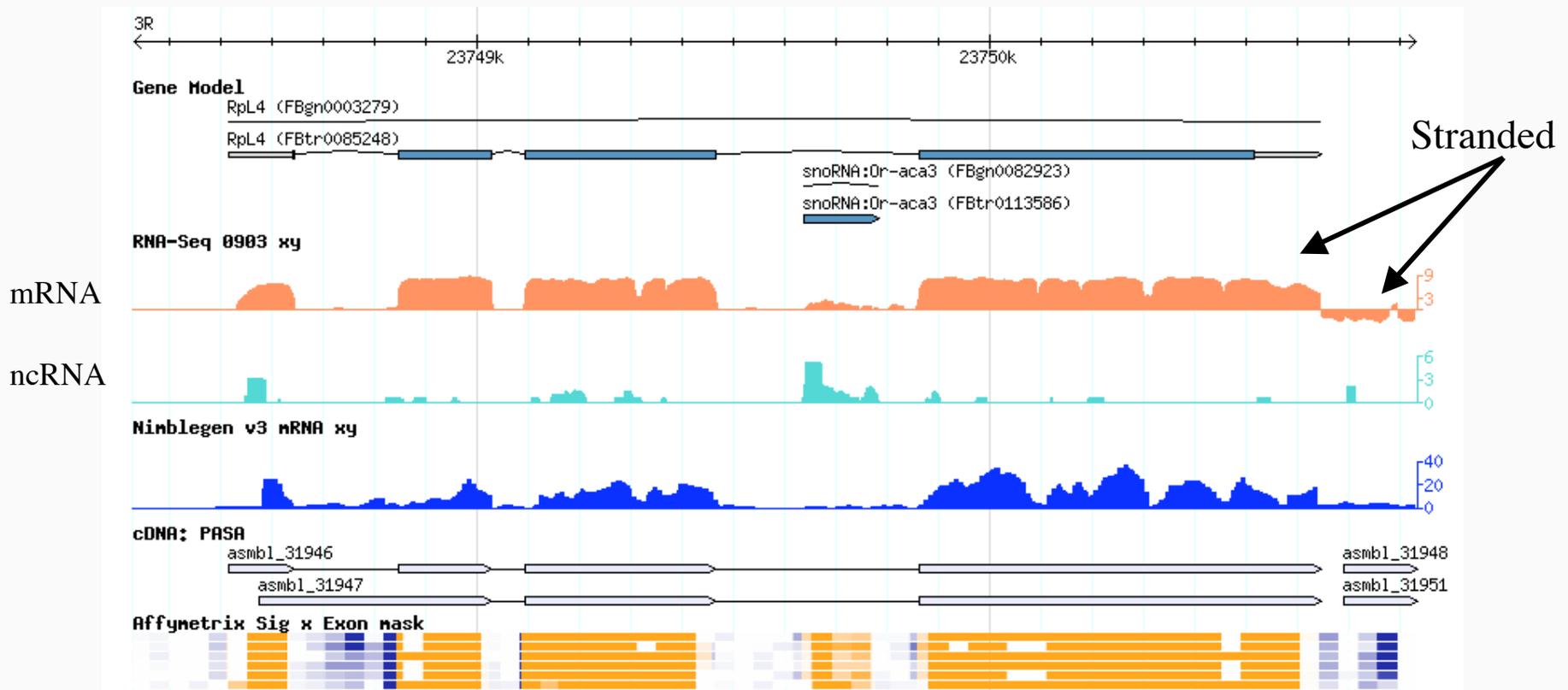
2008: Precision improves

Nimblegen has higher base precision than Affymetrix



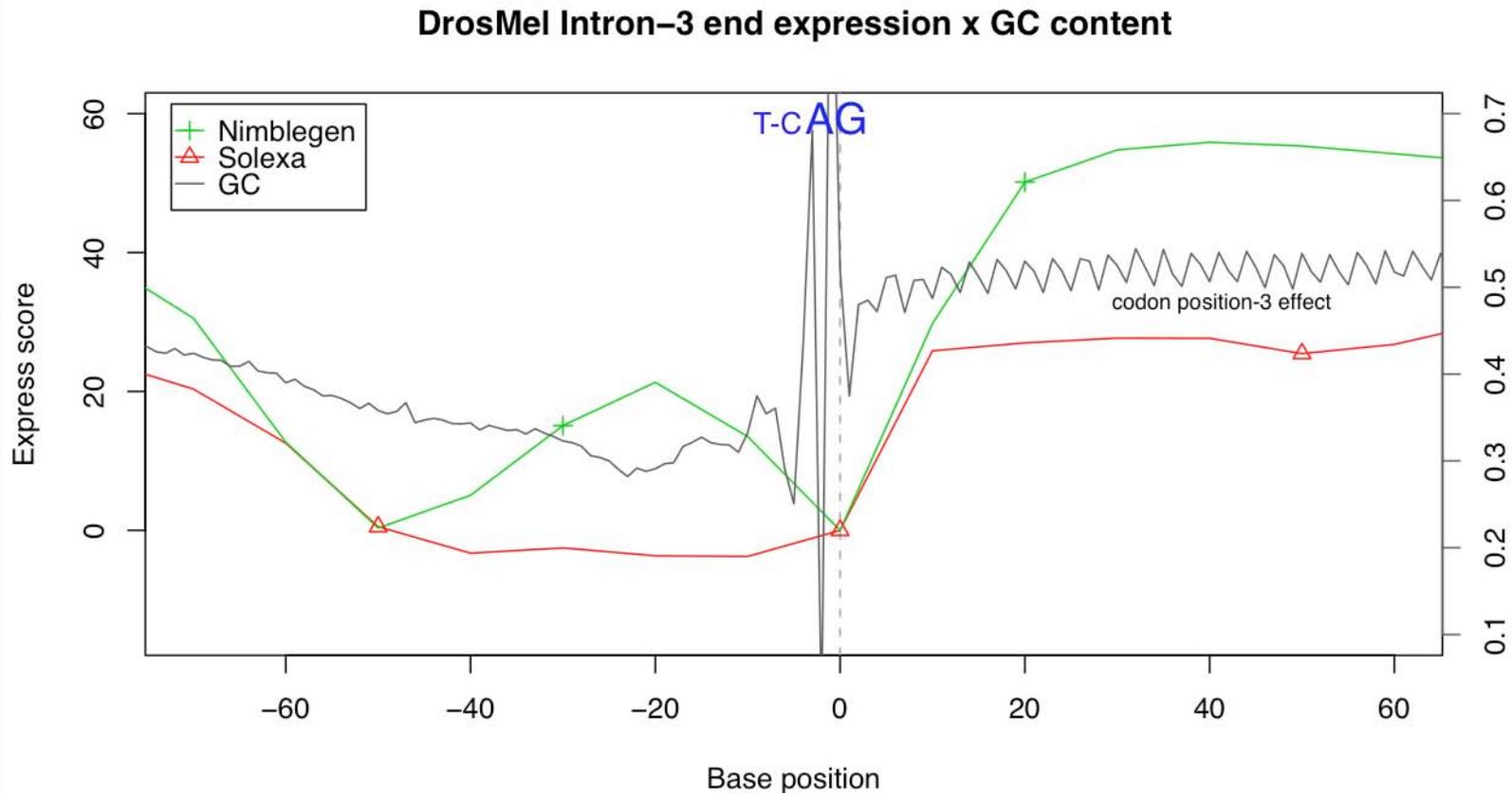
2009: Precision improves

RNA-Seq has higher base precision yet



And RNA-Seq here is strand-specific expression

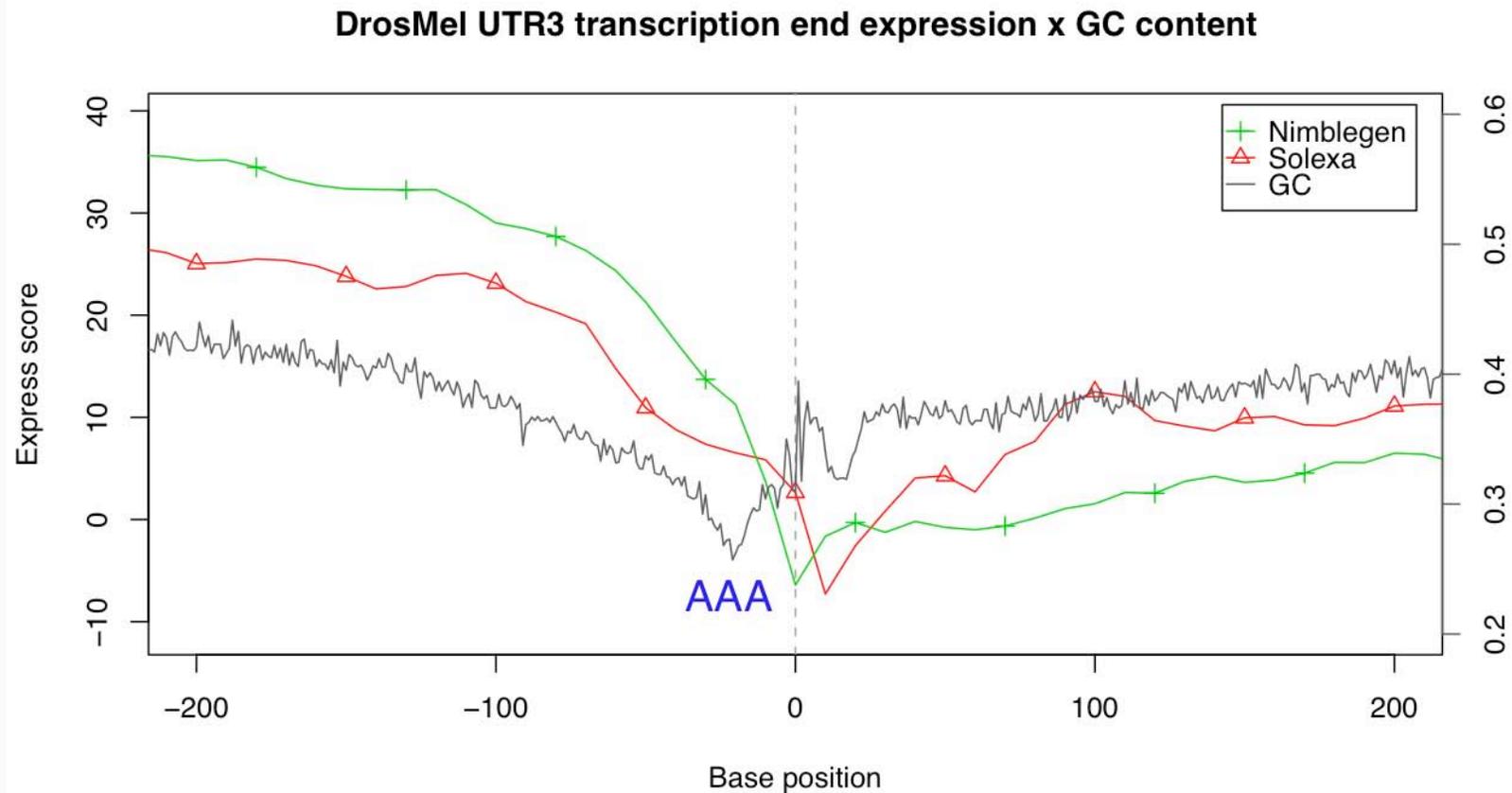
Intron-Exon Detection



<http://wfleabase.org/genome-summaries/tile-expression/tile-genestruct09/>

wfleabase.org/docs/tilexseq0904.pdf

Gene End Detection

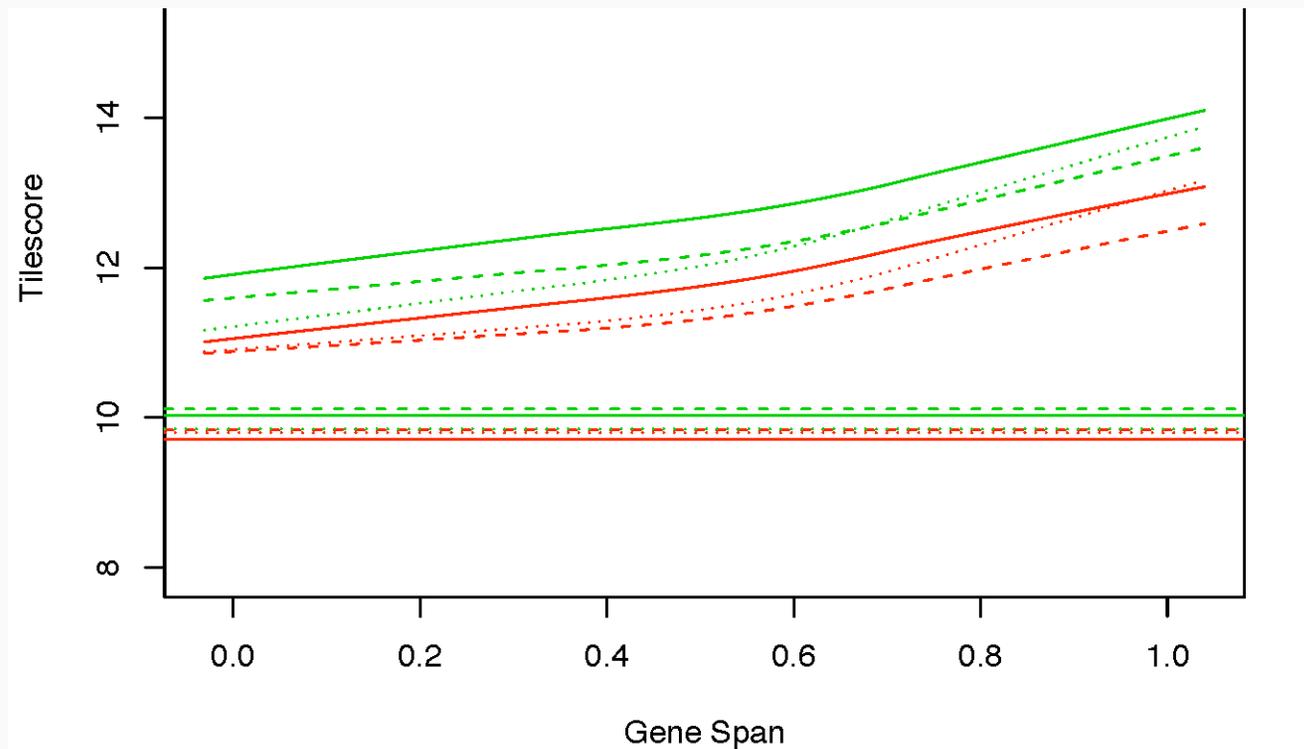


<http://wfleabase.org/genome-summaries/tile-expression/tile-genestruct09/>

wfleabase.org/docs/tilexseq0904.pdf

Differential expression

Gene end (3') has more expression, but

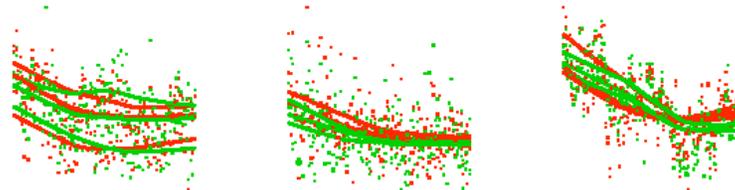


constant differential over gene span, *on average*

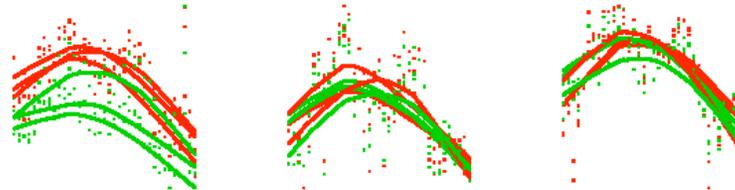
Differential expression

Genes vary in DE pattern over gene span

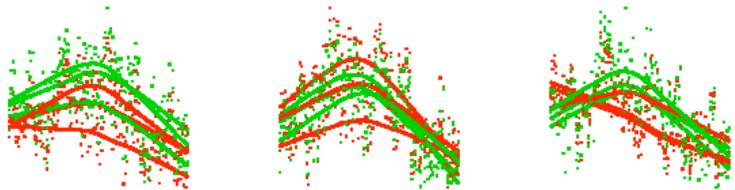
JGI_V11_208341
ntile: 91



JGI_V11_206450
ntile: 37



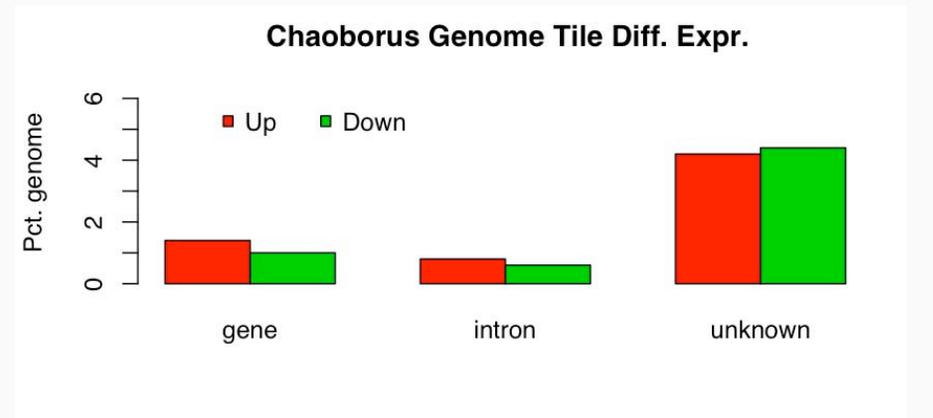
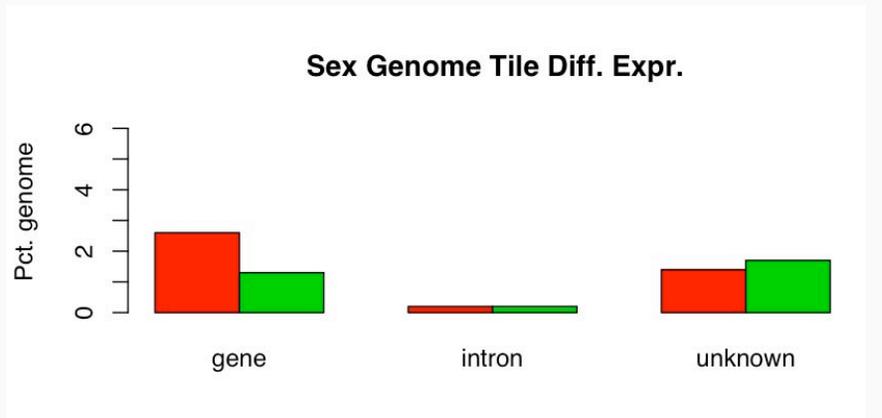
JGI_V11_186952
ntile: 101



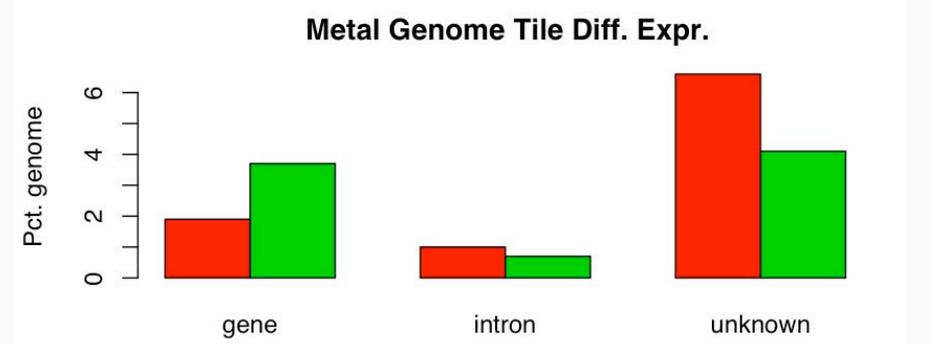
i.e., a few parts may not capture full gene expression

Differential expression

Gene models miss much expression

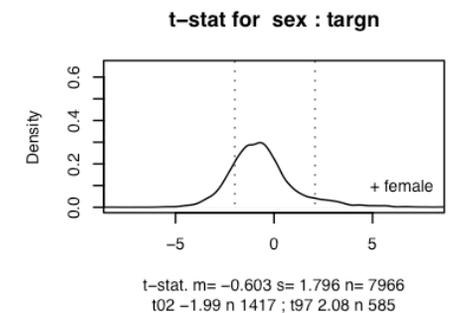
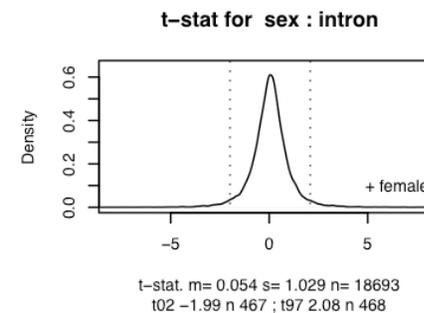
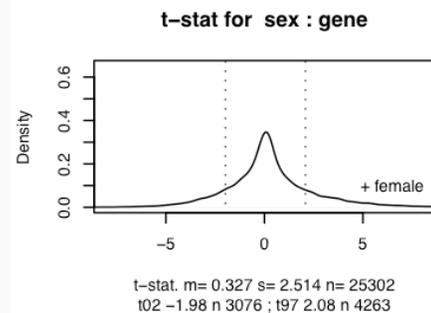
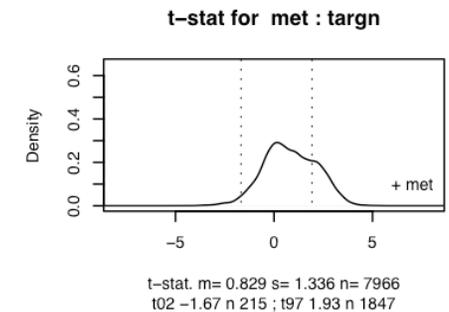
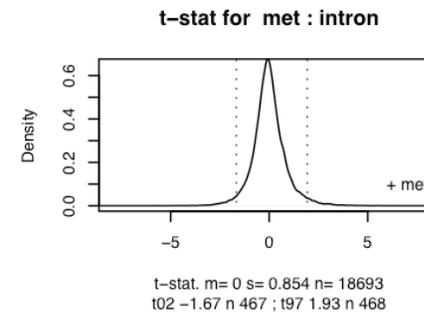
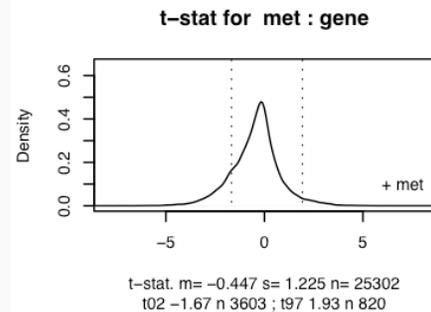
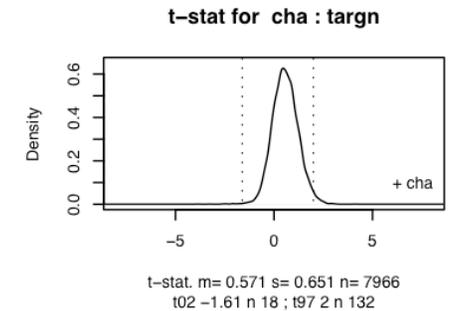
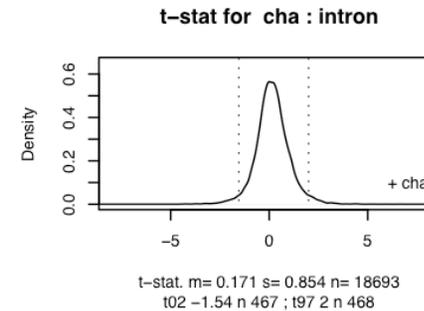
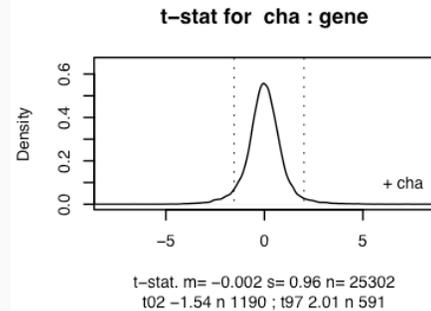


Known sex genes capture DE, but unknown regions capture environmental stress expression, in *Daphnia*.



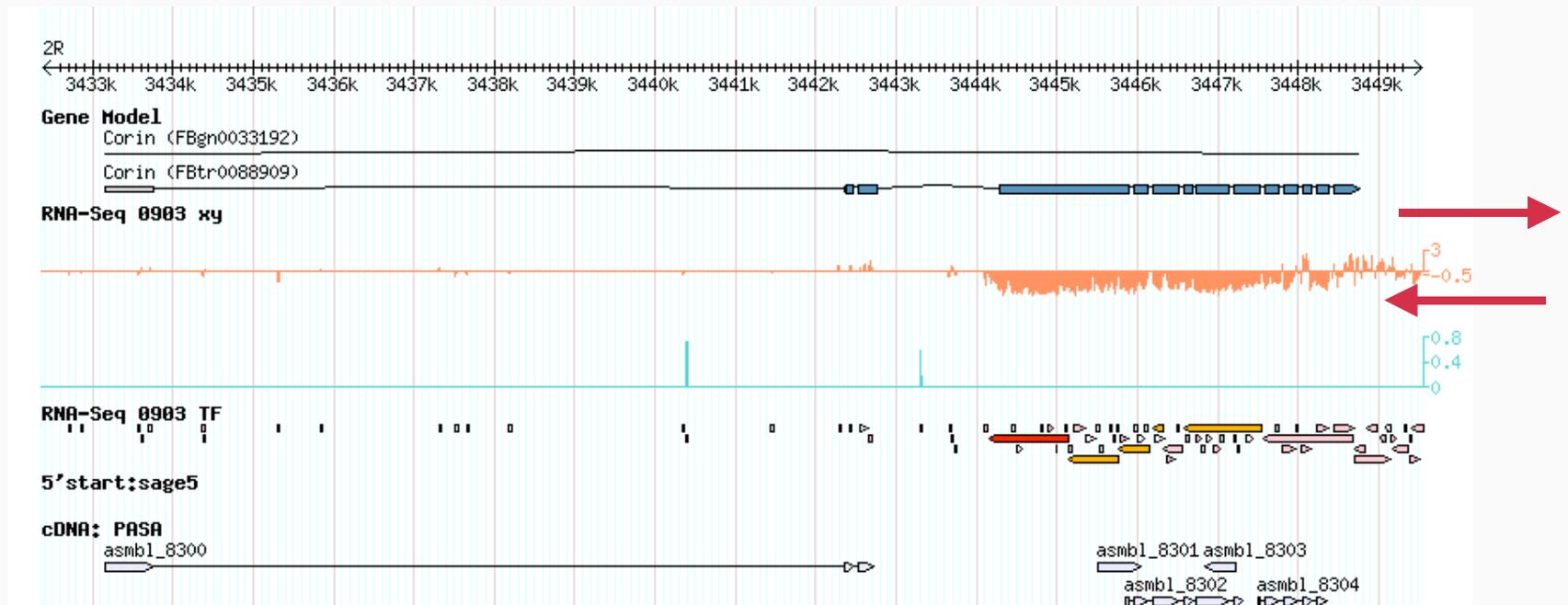
Differential eXp distribution

Introns show a null DE distribution, genes and TAR regions deviate from this. Use introns as baseline?



Stranded expression oddity

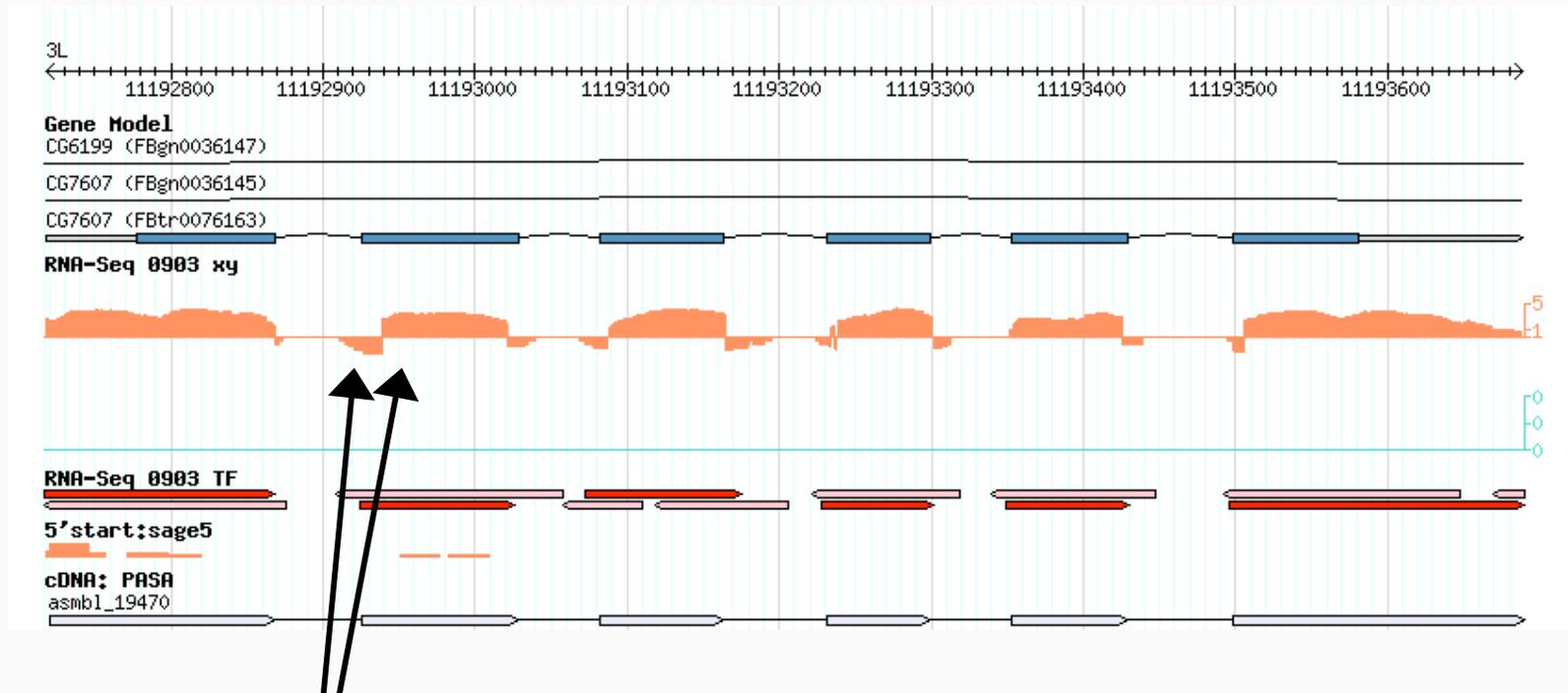
Expression in WRONG direction?



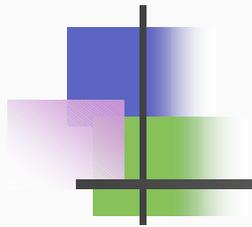
800 (3%) DrosMel Coding exons are expressed in wrong direction?

Stranded expression oddity

DrosMel transcript stranded expression (mRNA-Seq) picks up some odd cases.

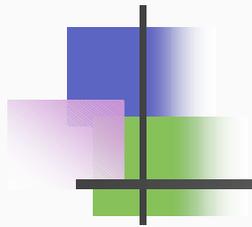


Expression in BOTH directions?



Summary

1. **Base-level expression (tiles, rna-seq) measures gene expression better**
 - Balances sensitivity (false rejection) with specificity (false discovery)
2. **Genome-wide base-level expression finds new genes**
 - As an alternative to gene-level studies, it has values and drawbacks. Computational methods need to improve to use this data well.
3. **Base-level expression measures gene structures well**
 - On average, and precision is improving for individual genes.
 - Gene oddities are found also



End note

Summary pages

wfleabase.org/genome-summaries/tile-expression/

insects.eugenes.org/species/data/dmel5/modencode/

Genome expression maps

insects.eugenes.org:8091/gbrowse/cgi-bin/gbrowse/drosmelme/

- expression in 52 cell lines (affy) and more precise solexa & nimblegen for a few cell lines

insects.eugenes.org:8091/gbrowse/cgi-bin/gbrowse/daphnia_pulex8/

- expression among 4 treatment groups (sex, metal stress, biotic predator); nimblegen

Gene Boundary Detection

