



# **Transcriptome Sequencing Reveals Alternative Splicing Events in Chemotherapy Resistant Colon Cancer Cells**

**Malachi Griffith**

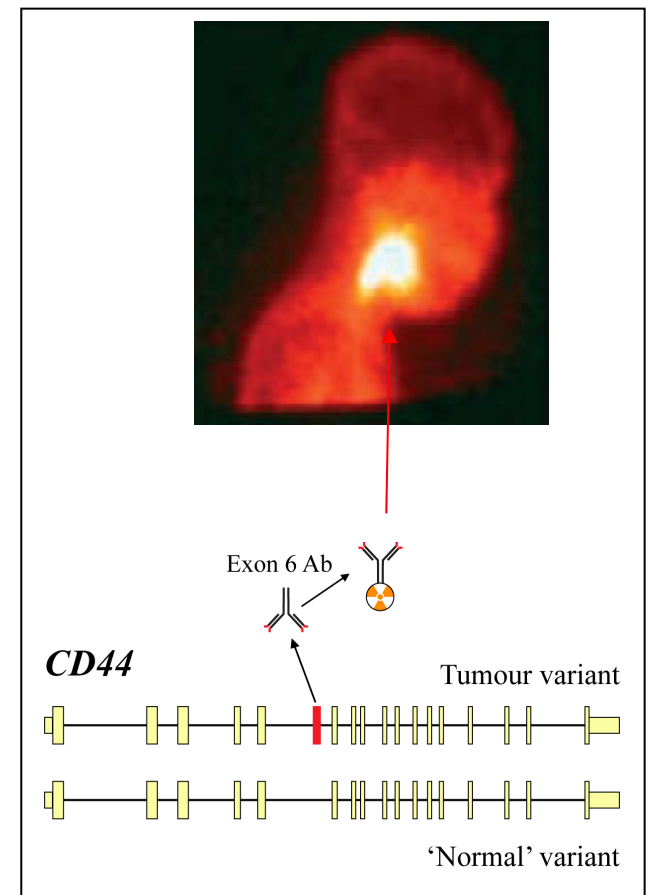
(Marco Marra lab)

AGBT February 5 2009



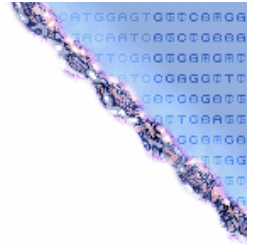
# Biomarker discovery is an important element of cancer research

- Biomarker discovery
  - Differential gene expression
  - Mutation detection
- Diagnosis, prognosis, treatment selection, therapeutic targets
  - e.g. *PSA*, *CEA*
- Specific transcript variants may represent useful biomarkers



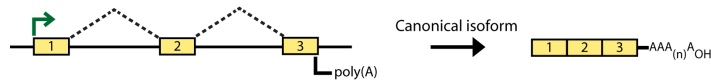
(Venables, 2006)

# Transcript variation is important to biomarker discovery

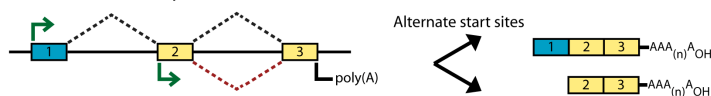


## Generates proteome diversity

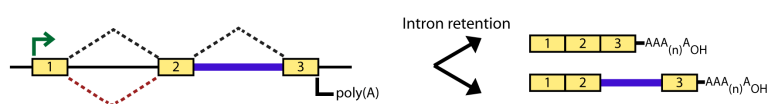
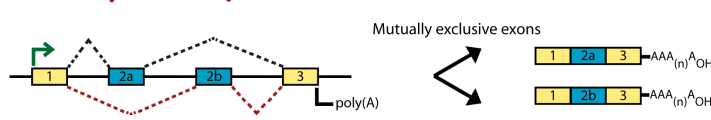
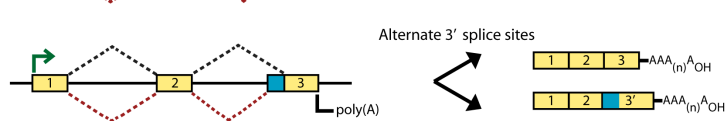
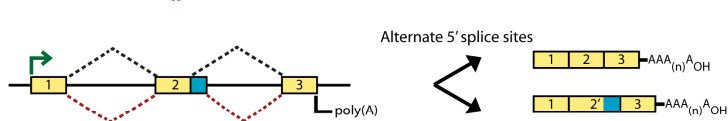
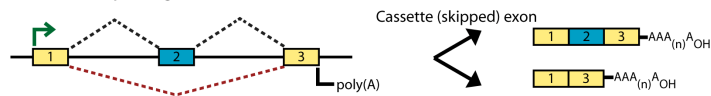
### Simple transcription



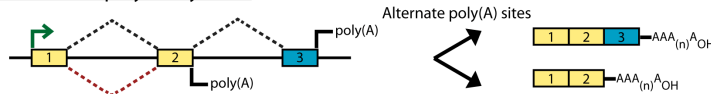
### Alternative transcript initiation



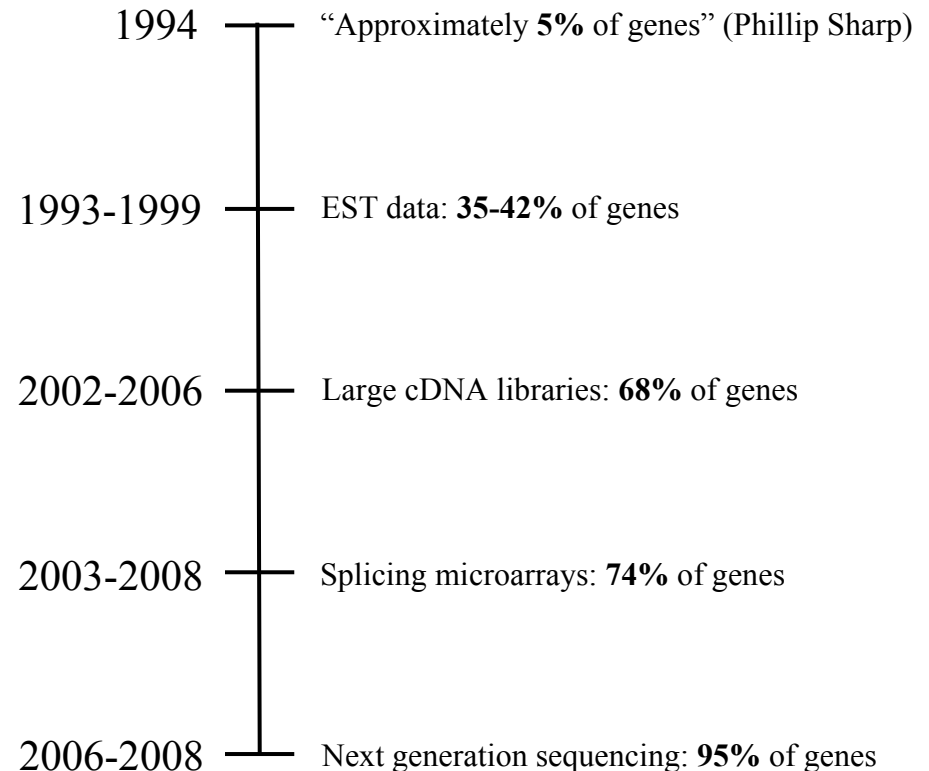
### Alternative splicing



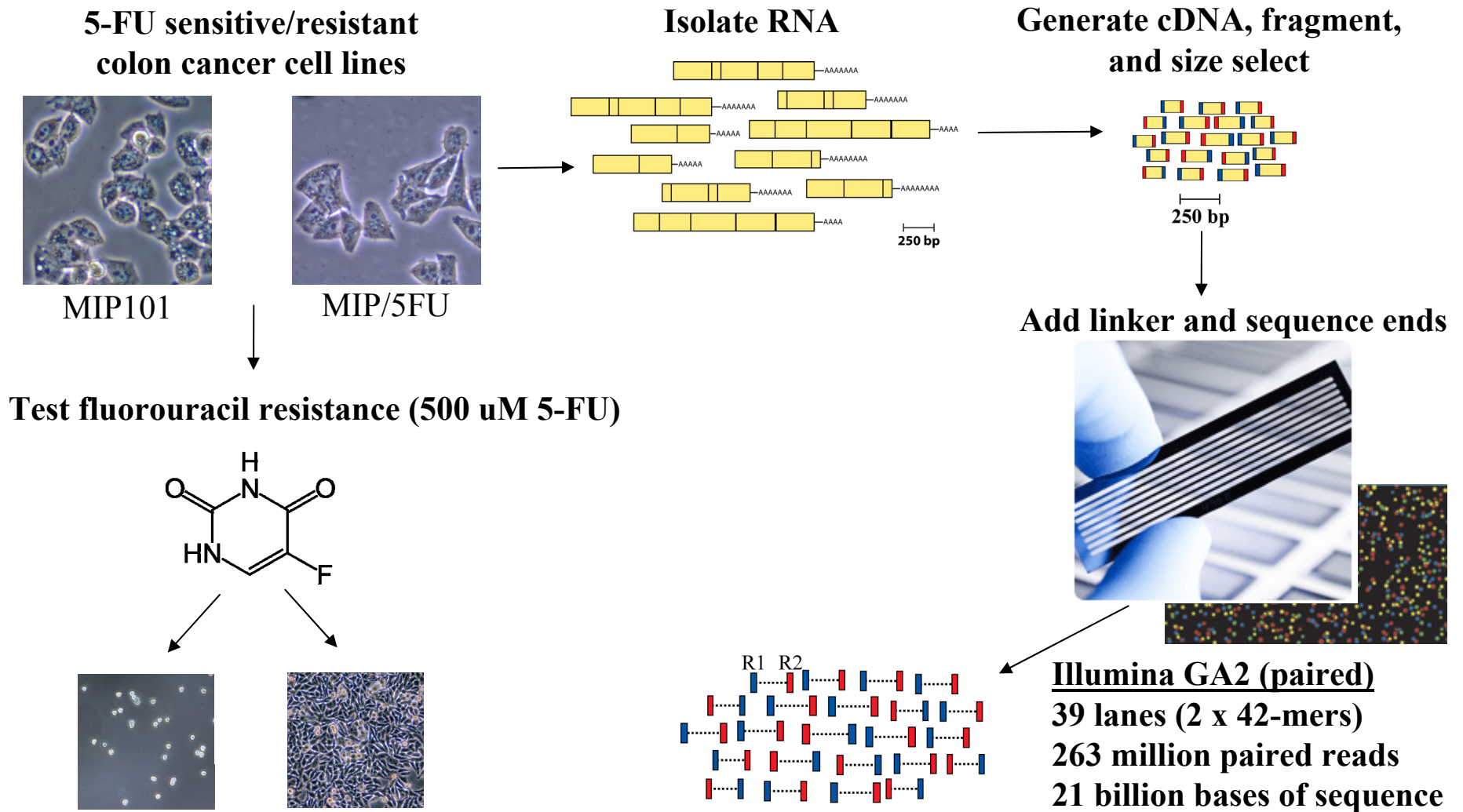
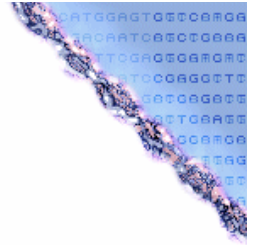
### Alternative polyadenylation



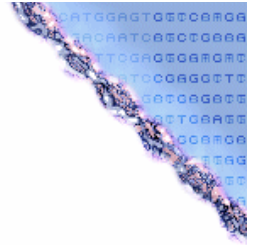
## Affects most genes



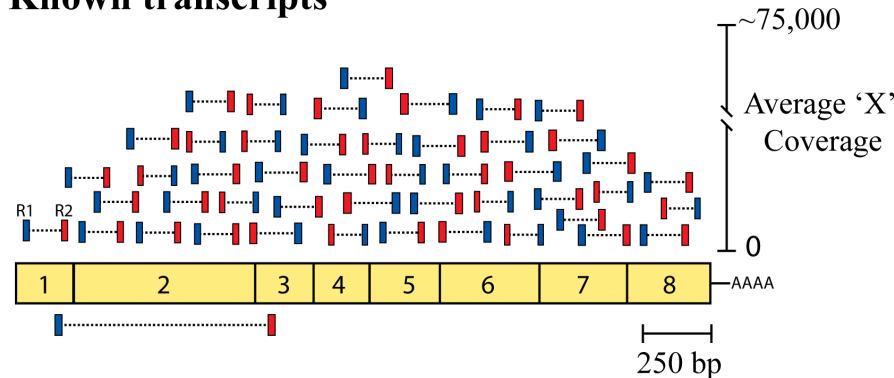
# 5-FU sensitive and resistant cancer cells were profiled by RNA sequencing



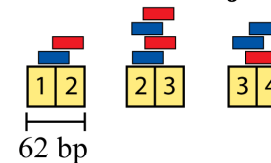
# Paired-end reads produced a nearly ‘complete’ transcriptome sequence



**Known transcripts**



**Known exon junctions**



**Predicted exon junctions**

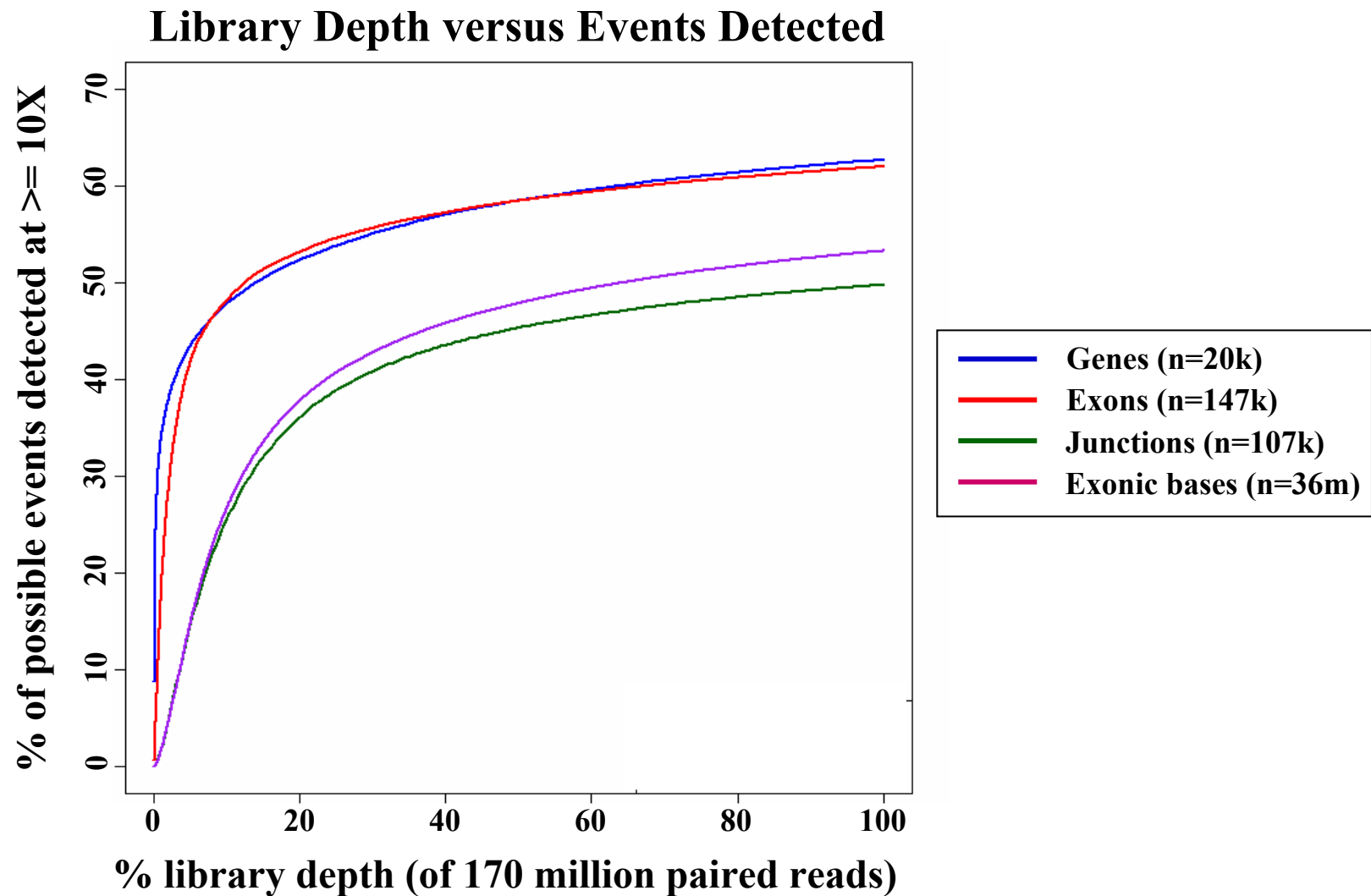
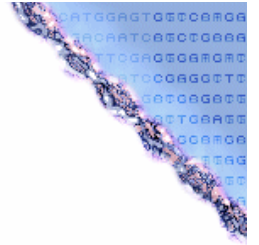


- 70% of all reads map to a known transcript or junction
- Detection by at least 10 reads:
  - 20k genes
  - 147k exons
  - 107k junctions
  - 36m exonic bases positions

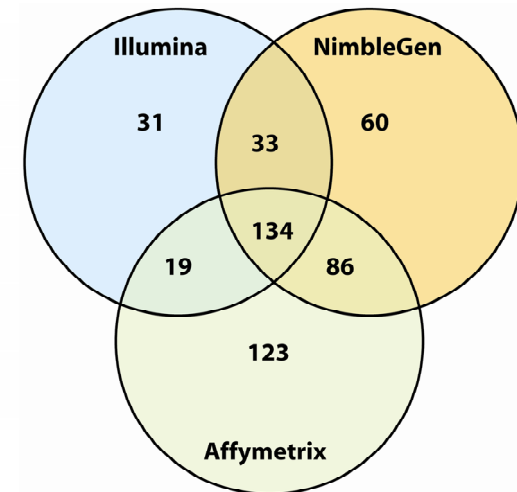
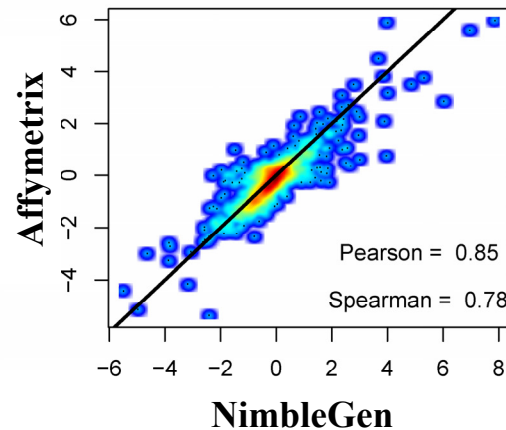
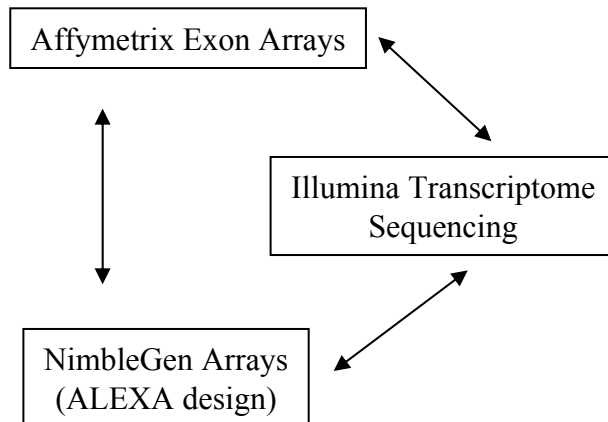
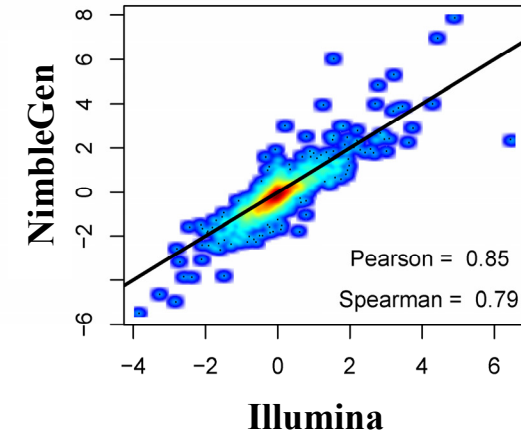
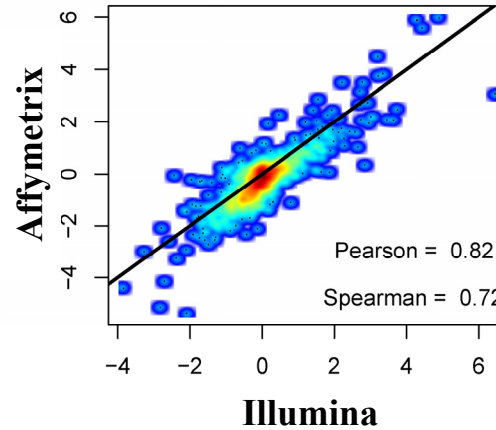
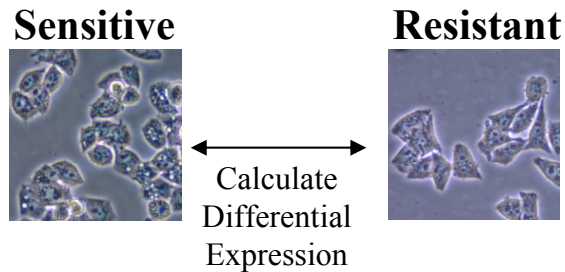




Detection of genes, exons, junctions and bases approached a plateau as library depth was increased



# Accurate differential expression estimates were derived (comparison to Affymetrix and NimbleGen arrays)

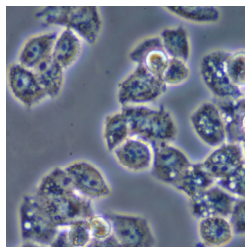


# Differential expression (DE) of genes and exons was readily detected

- 32 genes and 299 exons exhibit  $> 4$ -fold DE
- Bias towards loss of expression in resistant cells

Example: *H19*, a maternally imprinted non-coding gene

**Sensitive**



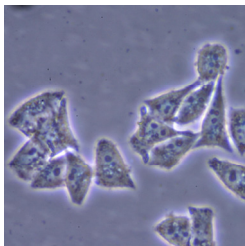
***H19***



~2.2 million reads (rank #1)

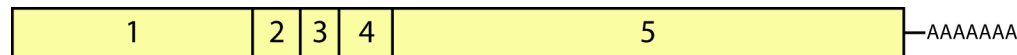
~21,000x coverage

**Resistant**



~90-fold down-regulated

***H19***



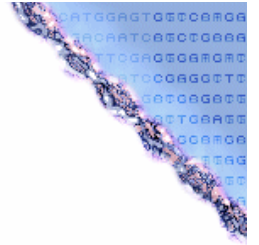
13,609 reads (rank #2083)

~229x coverage





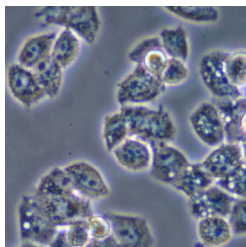
# Differential expression of alternatively transcribed isoforms was also apparent



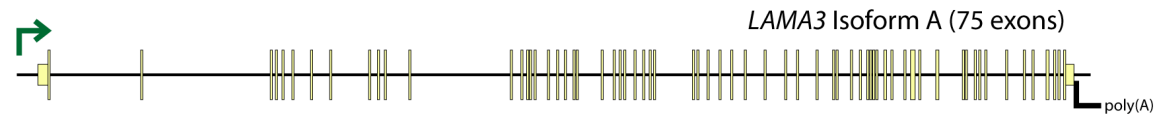
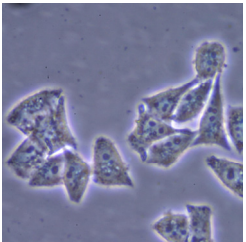
- ~20 genes show evidence of alternative transcript initiation or poly-adenylation

Example: *LAMA3*, laminin, alpha 3

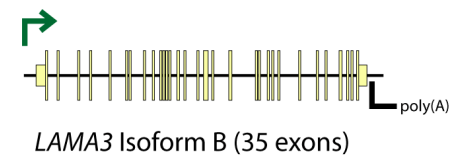
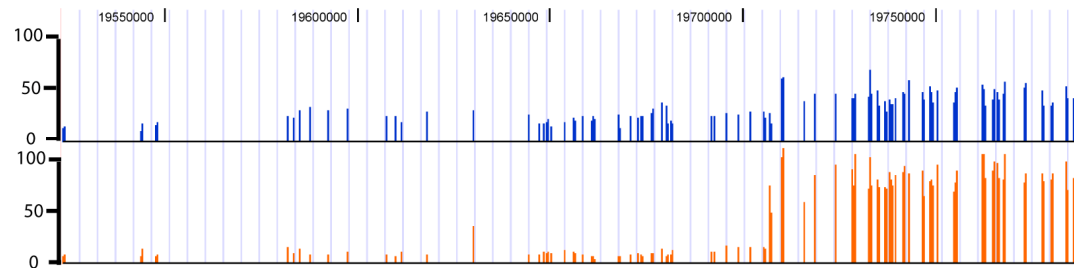
Sensitive



Resistant



Observed base coverage

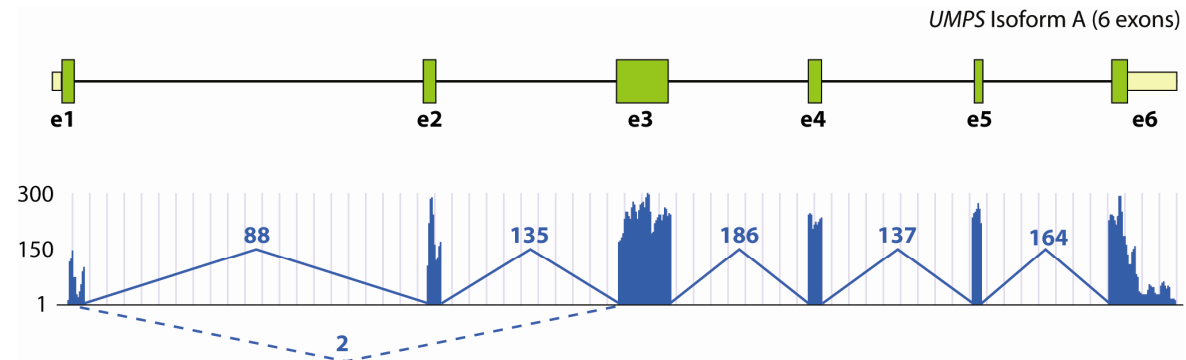
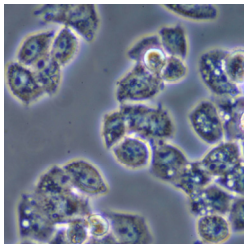


# Differential expression of exon-skipping isoforms was also apparent

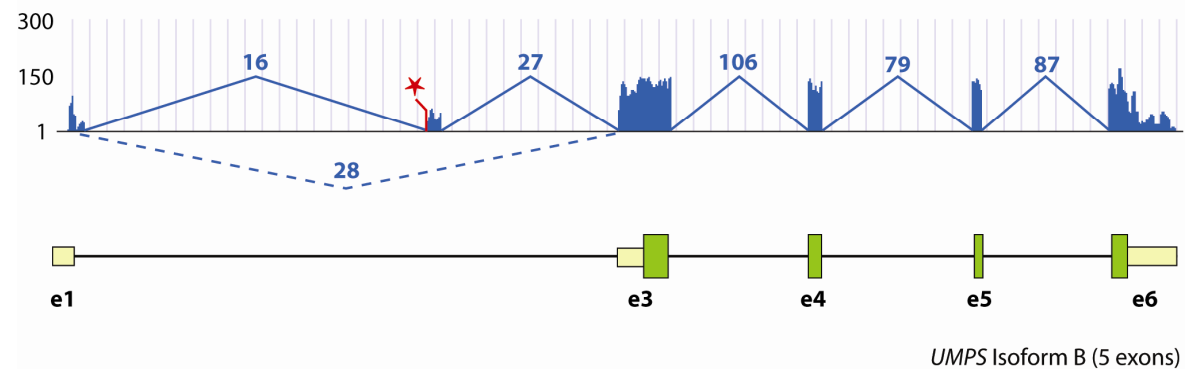
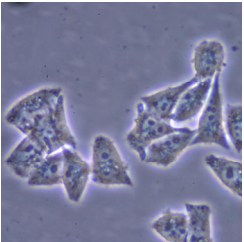
- 49 genes show evidence of differential exon skipping... 35 known and 14 novel.

Example: *UMPS*, uridine monophosphate synthetase

**Sensitive**

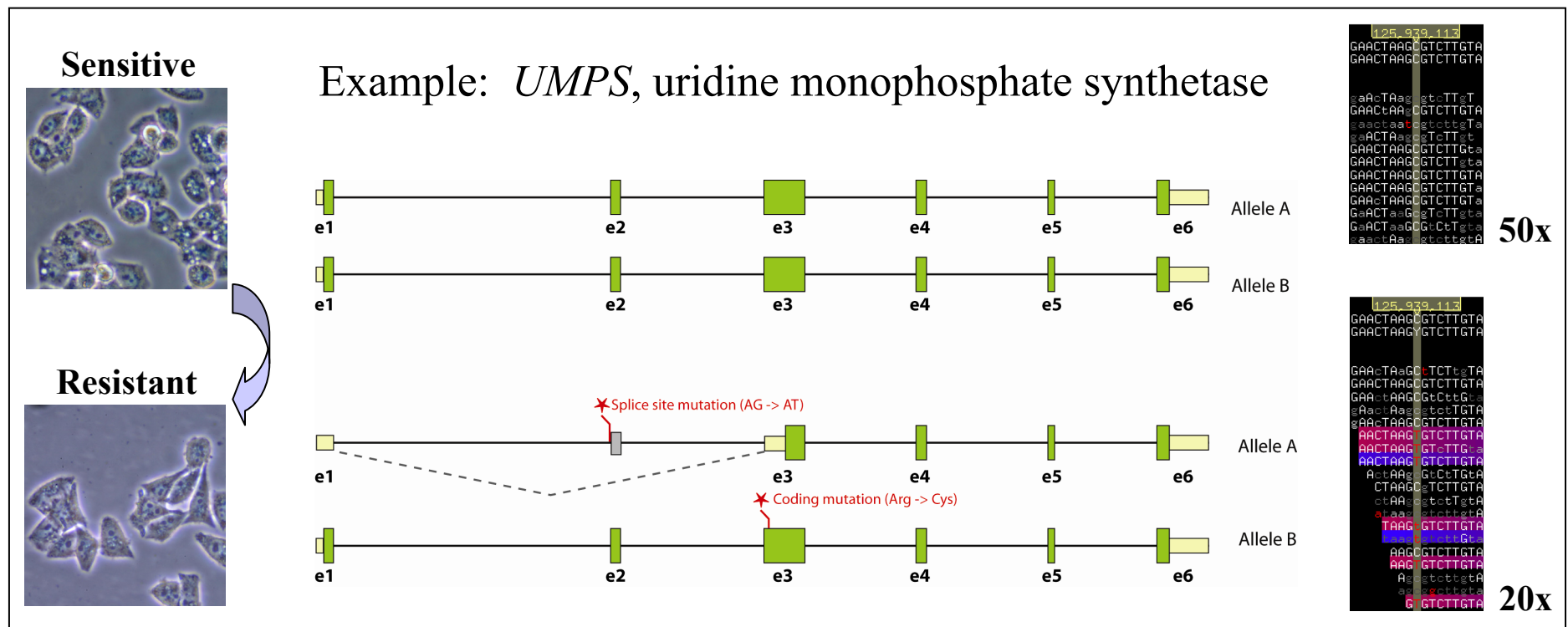


**Resistant**



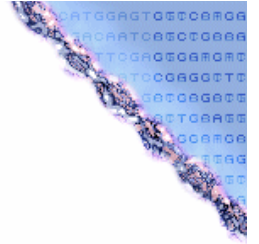
# Coding mutations acquired in resistant cells were also identified

- 20-50 genes show evidence of coding mutations acquired in resistant cells





# Conclusions

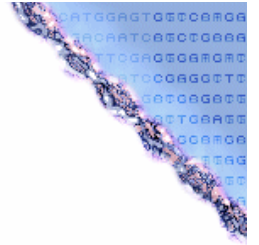


- Massively parallel sequencing allows comprehensive transcriptome re-sequencing
- Allows detection of differences between conditions beyond RNA expression levels
- 5-FU resistance may occur by redundant disruption of drug transport, metabolism and apoptosis pathways
- Clinical validation of *UMPS*
  - cohort of 100 pre- and post-treatment samples





# Acknowledgements



## **Supervisor**

Marco Marra

## **Committee**

Joseph Connors  
Stephane Flibotte  
Steve Jones  
Gregg Morin

## **Bioinformatics**

Obi Griffith  
Anca Petrescu  
Ryan Morin  
Allen Delaney

## **Laboratory**

Trevor Pugh  
Tesa Severson

## **Illumina**

Martin Hirst  
Thomas Zeng  
Yongjun Zhao  
Helen McDonald

## **Affymetrix**

Agnes Baross  
Susanna Chan  
Jennifer Asano  
Adrian Ally

## **Students**

Jessica Paul  
Lisa Miao

## **GI cancer**

Isabella Tai  
Michelle Tang

## **Clinical**

Sharlene Gill  
David Owen  
Margaret Luk  
Carl Brown  
Riva Pickering

## **Technology Development**

Patrick Rebstein



[illegible]

# Observed distribution of fragment sizes and relative transcript position

