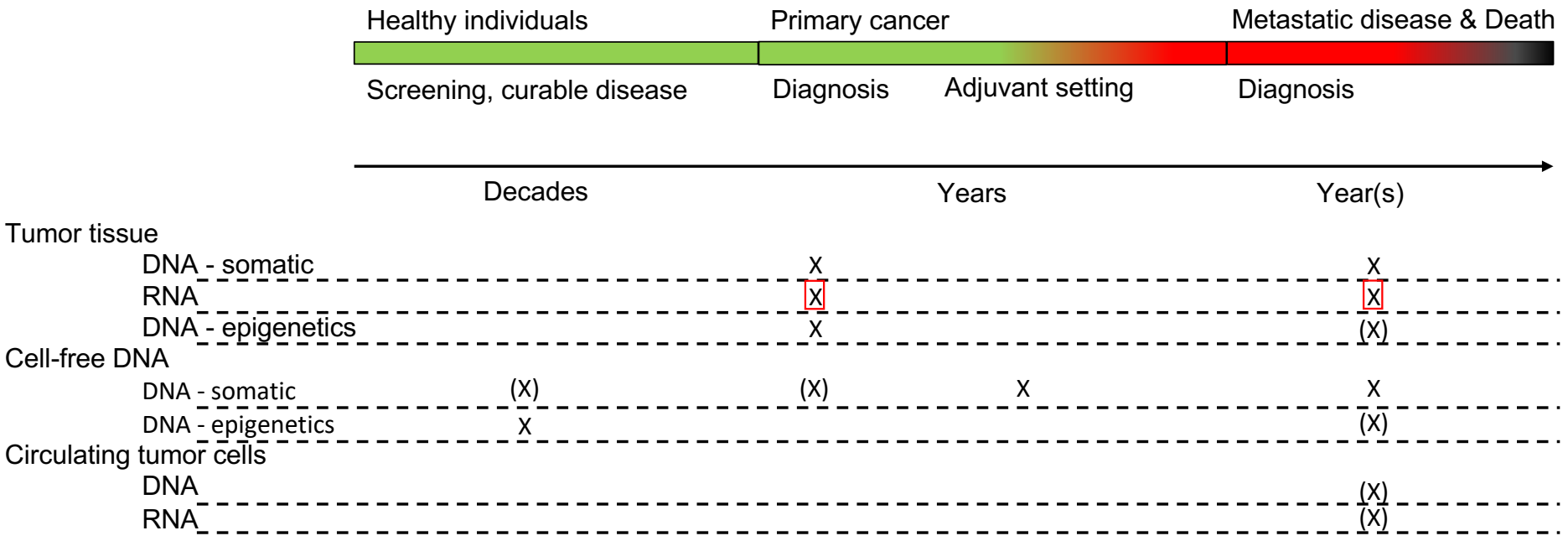
RNA-sequencing

Karolinska Institutet

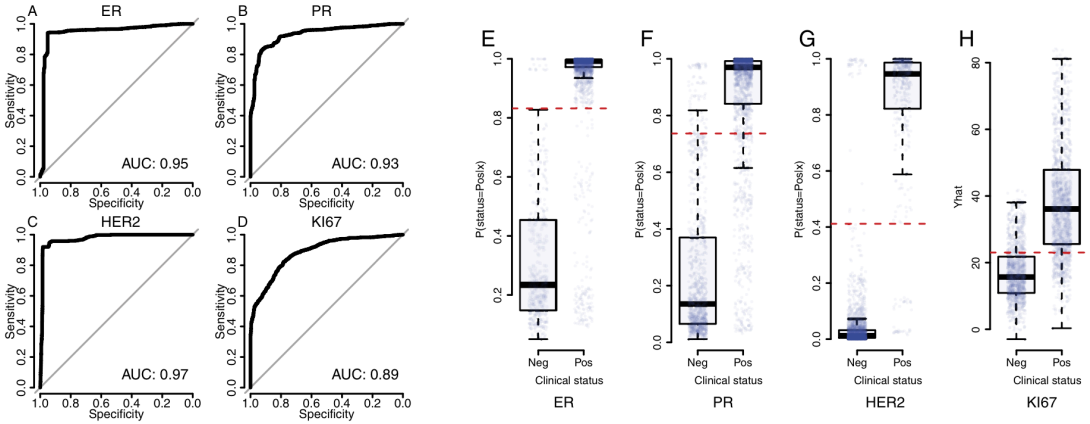
Tissue, analyte and context …



RNA-seq, epigenetic analysis and diagnostic relevance

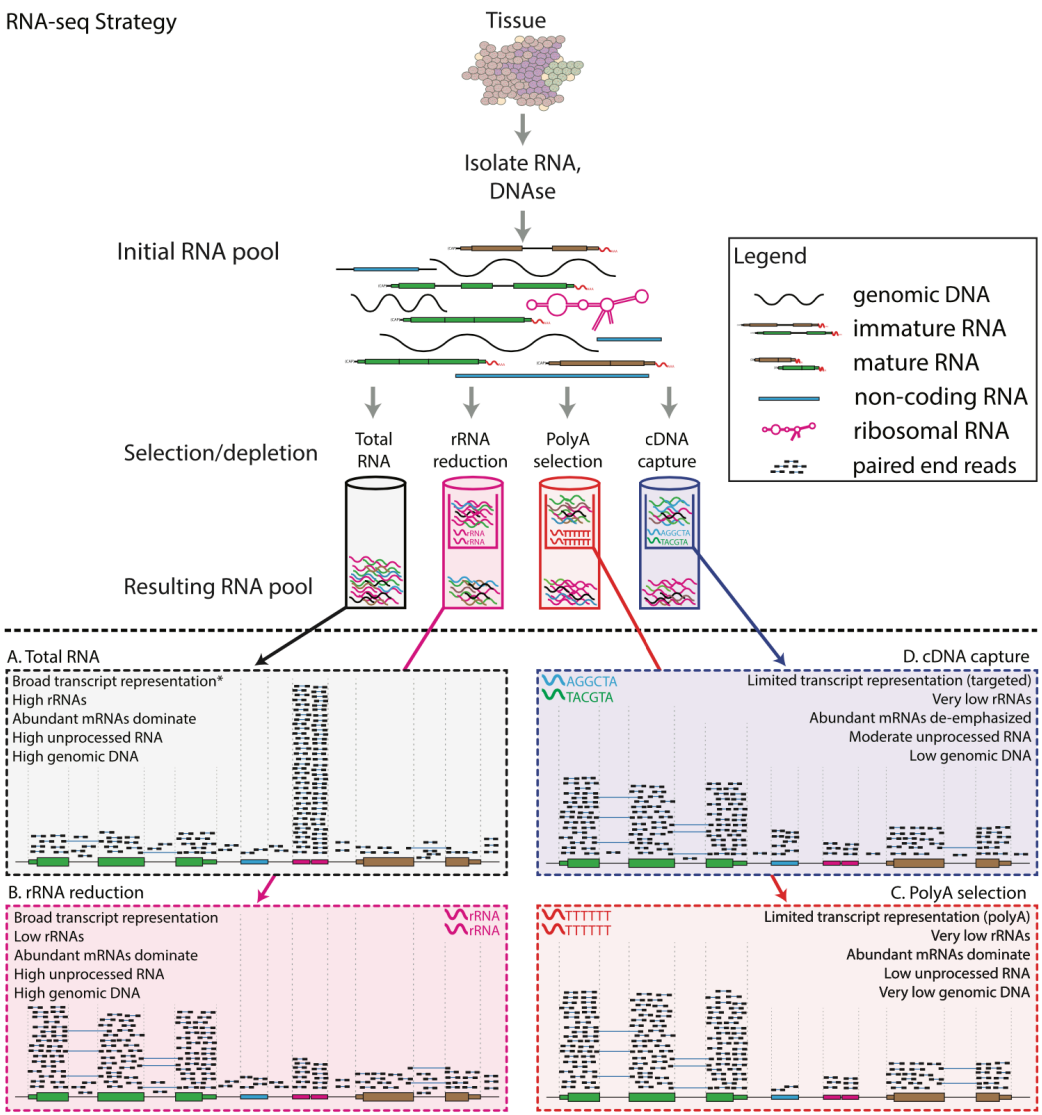
- RNA-seq in cancer
  - High dimensional data
    - Classification
    - Prognostication
    - Prediction modelling
  - Research
    - Splicing
    - Neoantigen expression
    - Differential gene expression
  - Direct diagnostic relevance
    - Outlier kinase expression
    - Fusion calling
    - Classification of cancer of unknown primary

RNA-seq based classification of BC clinical biomarkers



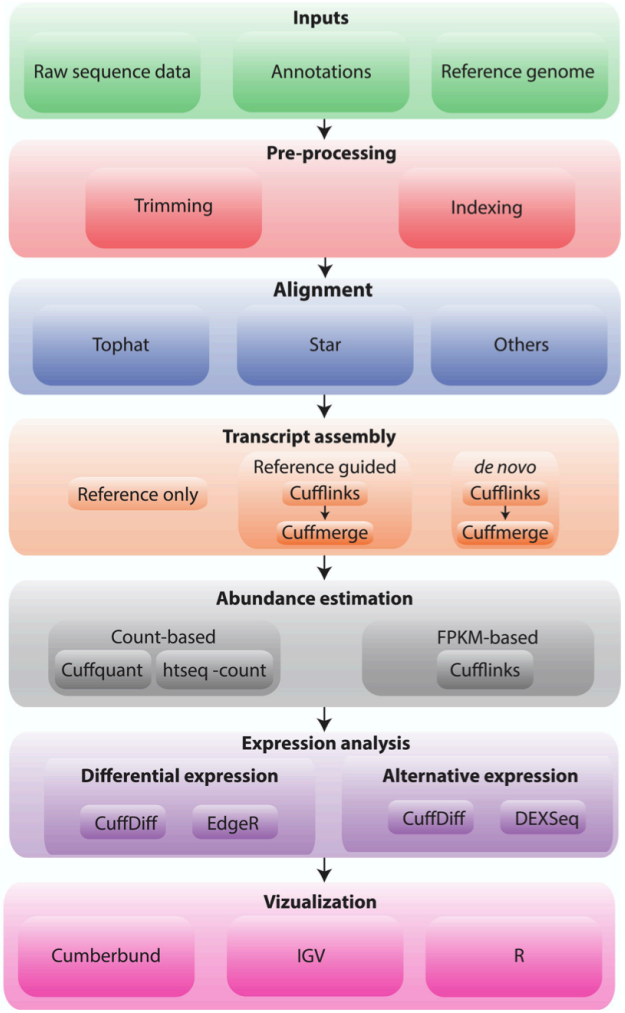
Sequencing-based breast cancer diagnostics as an alternative to routine biomarkers, Sci Rep 2016

RNA-seq strategies



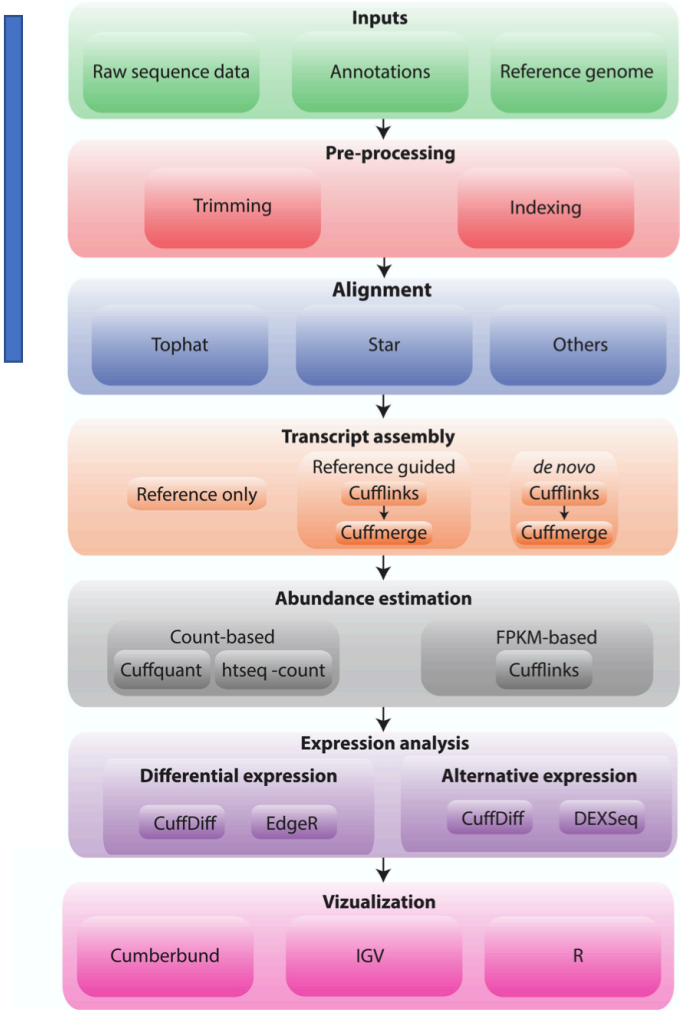
Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud, PLOS Computational Biology 2015

RNA-seq analysis pipeline



Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud,
PLOS Computational Biology 2015

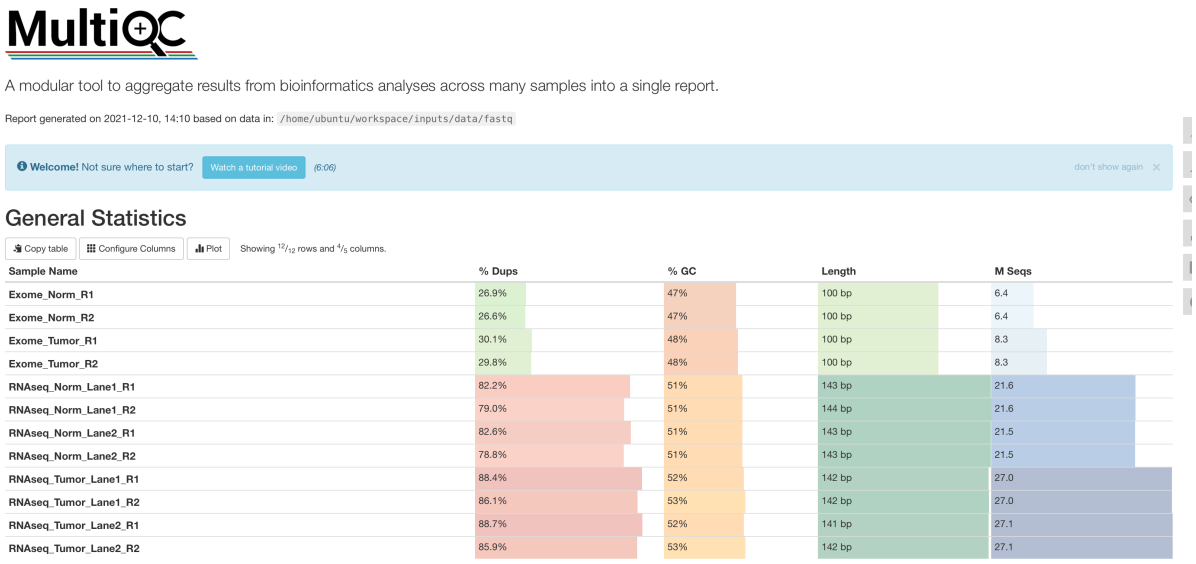
RNA-seq analysis pipeline



Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud,
PLOS Computational Biology 2015

RNA-seq, QC

- MultiQC will summarize
  - Adapter contamination
  - Ribosomal RNA fraction
  - Problems with library length
  - Fraction aligned reads etc
- Power of QC is having a background distribution



MultiQC

A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.

Report generated on 2021-12-10, 14:10 based on data in: `/home/ubuntu/workspace/inputs/data/fastq`

Welcome! Not sure where to start? Watch a tutorial video *(6:06)* don't show again

General Statistics

Copy table | Configure Columns | Plot | Showing 12/12 rows and 4/5 columns.

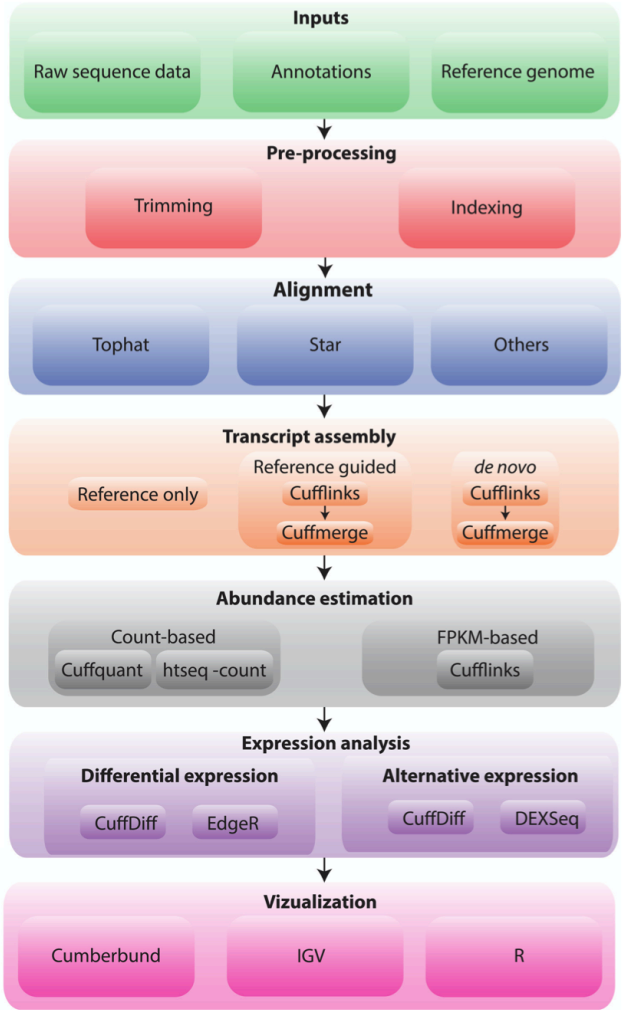
Sample Name	% Dups	% GC	Length	M Seqs
Exome_Norm_R1	26.9%	47%	100 bp	6.4
Exome_Norm_R2	26.6%	47%	100 bp	6.4
Exome_Tumor_R1	30.1%	48%	100 bp	8.3
Exome_Tumor_R2	29.8%	48%	100 bp	8.3
RNAseq_Norm_Lane1_R1	82.2%	51%	143 bp	21.6
RNAseq_Norm_Lane1_R2	79.0%	51%	144 bp	21.6
RNAseq_Norm_Lane2_R1	82.6%	51%	143 bp	21.5
RNAseq_Norm_Lane2_R2	78.8%	51%	143 bp	21.5
RNAseq_Tumor_Lane1_R1	88.4%	52%	142 bp	27.0
RNAseq_Tumor_Lane1_R2	86.1%	53%	142 bp	27.0
RNAseq_Tumor_Lane2_R1	88.7%	52%	141 bp	27.1
RNAseq_Tumor_Lane2_R2	85.9%	53%	142 bp	27.1

Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud,
PLOS Computational Biology 2015

RNA-seq analysis pipeline

In the course you will test:

- Kallisto
  - Reference only analysis
- HISAT
  - Transcript guided analysis



Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud,
PLOS Computational Biology 2015

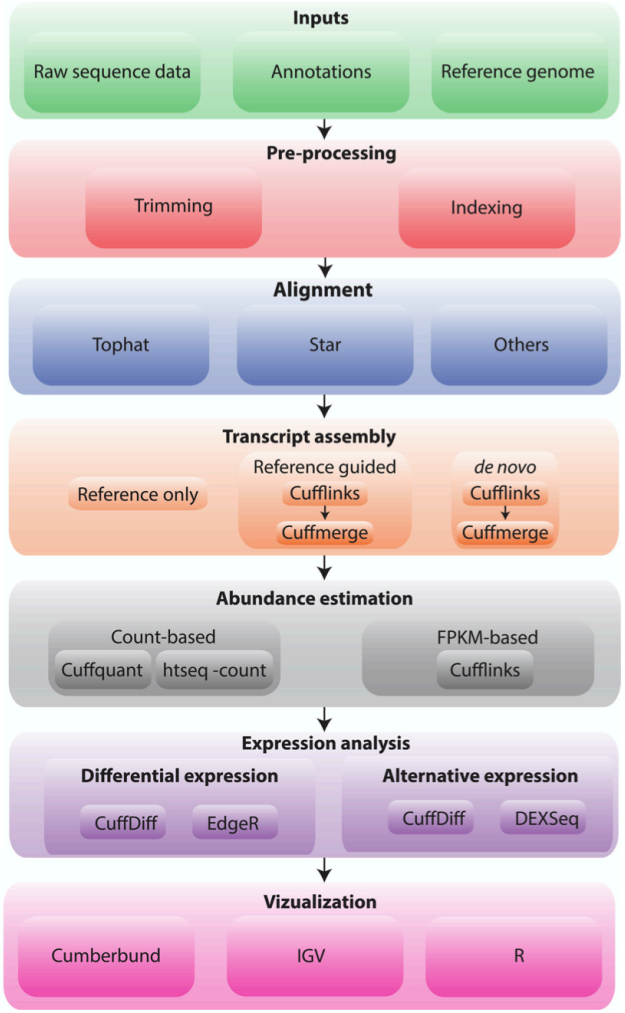
RNA-seq analysis pipeline

Expression levels vary a lot (10^5 – 10^7 times)

We will use TPM (Transcript per Kilobase Million) to assess expression:

1) Divide each gene/transcript fragment count by length of each gene/transcript in kilobases
   - Fragments per kilobase, FPK
2) Sum all FPK values for the sample and divide by 1,000,000
   - "per million" scaling factor
3) Divide #1 by #2 (TPM)

The sum of all TPMs in each sample is the same.



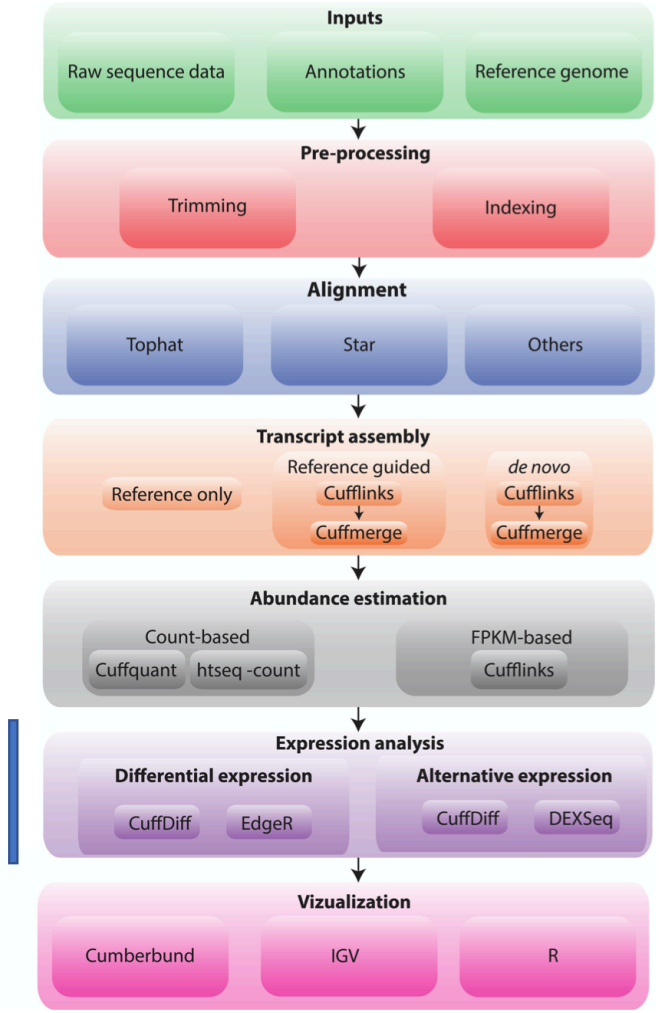
Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud, PLOS Computational Biology 2015

RNA-seq analysis pipeline

Normalization required.

Spike ins might help.

…



Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud,
PLOS Computational Biology 2015

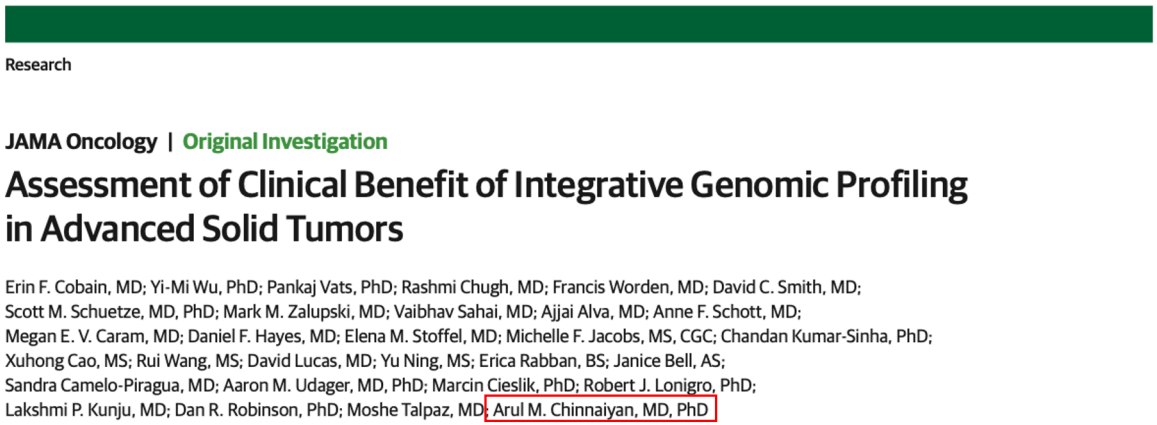
RNAseq paper from ICGC

- 1188 cases with RNAseq + WGS
- Copy-number alterations were the major drivers of variations in total gene and allele-specific expression.
- 82% of gene fusions were associated with structural variants

Genomic basis for RNA alterations in cancer

Clinical impact of comprehensive DNA and RNA sequencing

- The Michigan Oncology Sequencing Program
  - Tumor biopsy sequencing with paired gDNA
  - Whole-exome or targeted capture
  - RNA-sequencing
    - Fusion detection
    - Classification of Cancer Of Unknown Primary (CUP)
  - Inclusion of 1138 advanced/metastatic patients between 2011-2018
  - Clinical benefit rate from NGS-directed therapy



Research

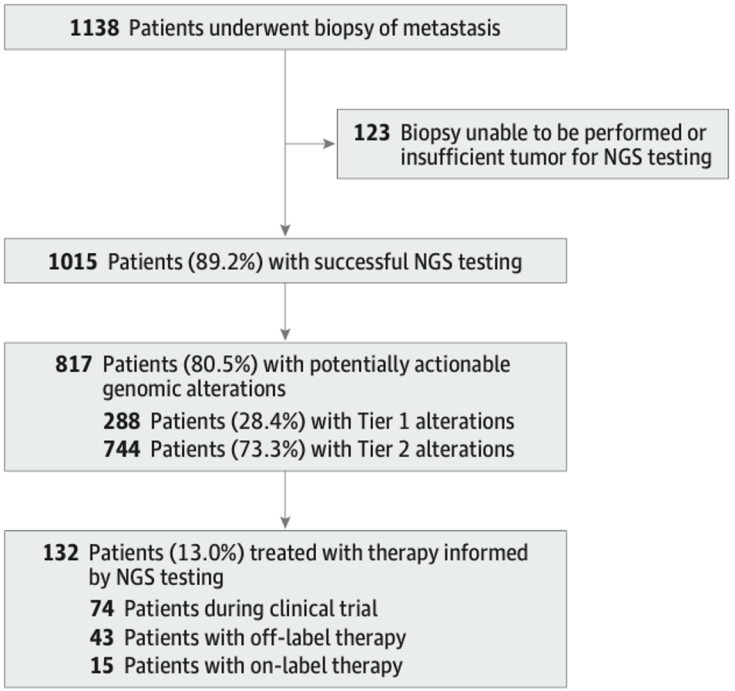
JAMA Oncology | Original Investigation

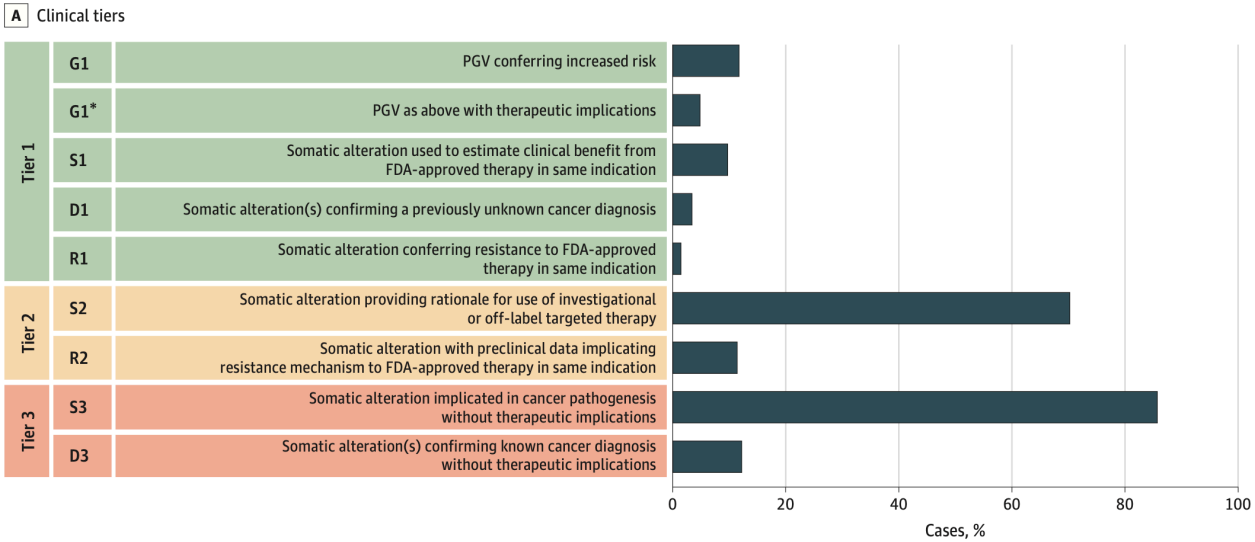
Assessment of Clinical Benefit of Integrative Genomic Profiling in Advanced Solid Tumors

Erin F. Cobain, MD; Yi-Mi Wu, PhD; Pankaj Vats, PhD; Rashmi Chugh, MD; Francis Worden, MD; David C. Smith, MD; Scott M. Schuetze, MD, PhD; Mark M. Zalupski, MD; Vaibhav Sahai, MD; Ajjai Alva, MD; Anne F. Schott, MD; Megan E. V. Caram, MD; Daniel F. Hayes, MD; Elena M. Stoffel, MD; Michelle F. Jacobs, MS, CGC; Chandan Kumar-Sinha, PhD; Xuhong Cao, MS; Rui Wang, MS; David Lucas, MD; Yu Ning, MS; Erica Rabban, BS; Janice Bell, AS; Sandra Camelo-Piragua, MD; Aaron M. Udager, MD, PhD; Marcin Cieslik, PhD; Robert J. Lonigro, PhD; Lakshmi P. Kunju, MD; Dan R. Robinson, PhD; Moshe Talpaz, MD; Arul M. Chinnaiyan, MD, PhD

What can we hope to achieve with iPCM/clinical implementation?

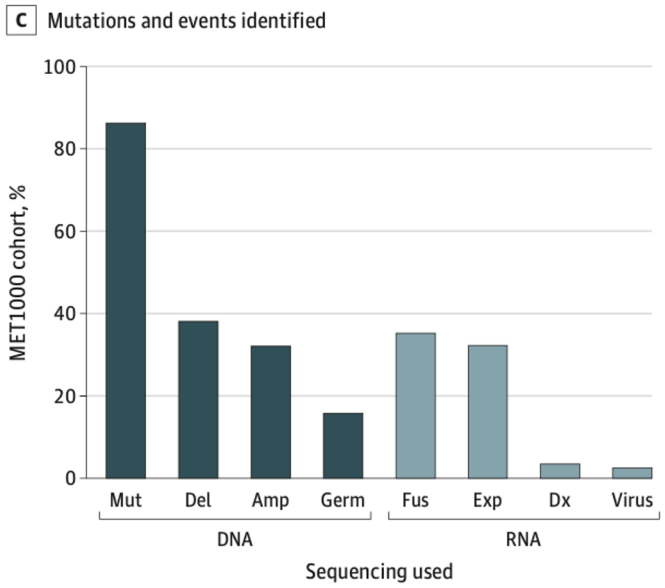
Figure 1. CONSORT Diagram of Patients in the MET1000 Cohort





What can we hope to achieve with iPCM/clinical implementation?

- 713 patients (70.2%) carried a potentially actionable somatic alteration
- 80.5% carried a clinically relevant alteration
  - 95% were identified by DNA sequencing
  - 63.5% were identified by RNA sequencing



- Sequencing-directed therapy was initiated in 132 patients
  - 49 experienced clinical benefit
  - 26 received therapy 12 months or longer
    - DDR-
    - MSI+
    - Gene fusion carriers
    - Hotspot mutations
    - Amplifications
- 169 pathogenic germline variants were detected
  - 155 were unknown
- 55 patients with cancer of unknown origin
  - 28 were re-classified

- The End