

SG-NEx, xPore, and bambu: Data and Tools for Nanopore RNA-Seq analysis

Chen Ying¹, Ploy Naruemon Pratanwanich¹, Laura Wratten¹, Yuk Kei, Christopher Hendra¹, Yao Fei¹, Andre Sim¹, Hwee Meng Low¹, Torsten Wüstefeld¹, Sarah Ng¹, Alexandre Thiery², Sho Goh¹, The Singapore Nanopore-Expression Consortium, [Jonathan Göke](mailto:Jonathan.Goekel@iglab.org)✉

¹Computational and Systems Biology, Genome Institute of Singapore
²Department of Statistics and Applied Probability, National University of Singapore

Follow @JonathanGoekel

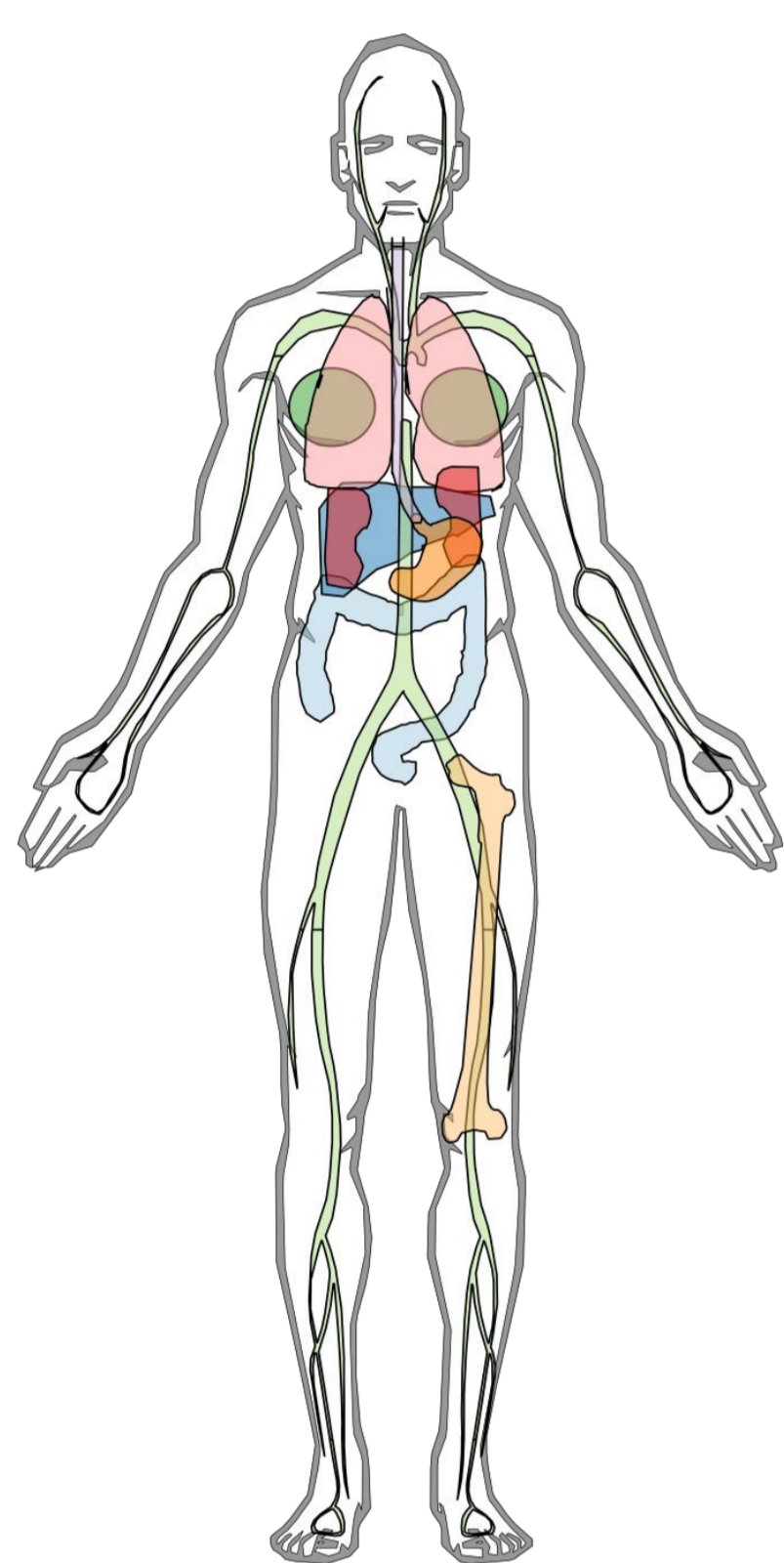
<https://github.com/goekelab> | www.iglab.org | gokej@gis.a-star.edu.sg



Summary

Long Read RNA-Sequencing enables a large number of applications. Here we present an overview of projects at the Genome Institute of Singapore aimed at generating data resources and methods to facilitate long read Nanopore RNA-Seq analysis.

SG-NEx: The Singapore Nanopore-Expression Project



SG-NEX

Liver	HepG2	direct RNA
Leukocytes	K562	direct cDNA
Lung	A549	PCR cDNA
Breast	MCF7	short read cDNA
Colon	Hct116	

The Singapore Nanopore Expression Project (SG-NEx) is a resource for systematically generated long read RNA-Sequencing data. The resource comprises 5 cell lines which were sequenced using direct RNA-Seq, direct cDNA-Seq, and cDNA-PCR Sequencing. Each cell line was sequenced with more than 30 million reads.

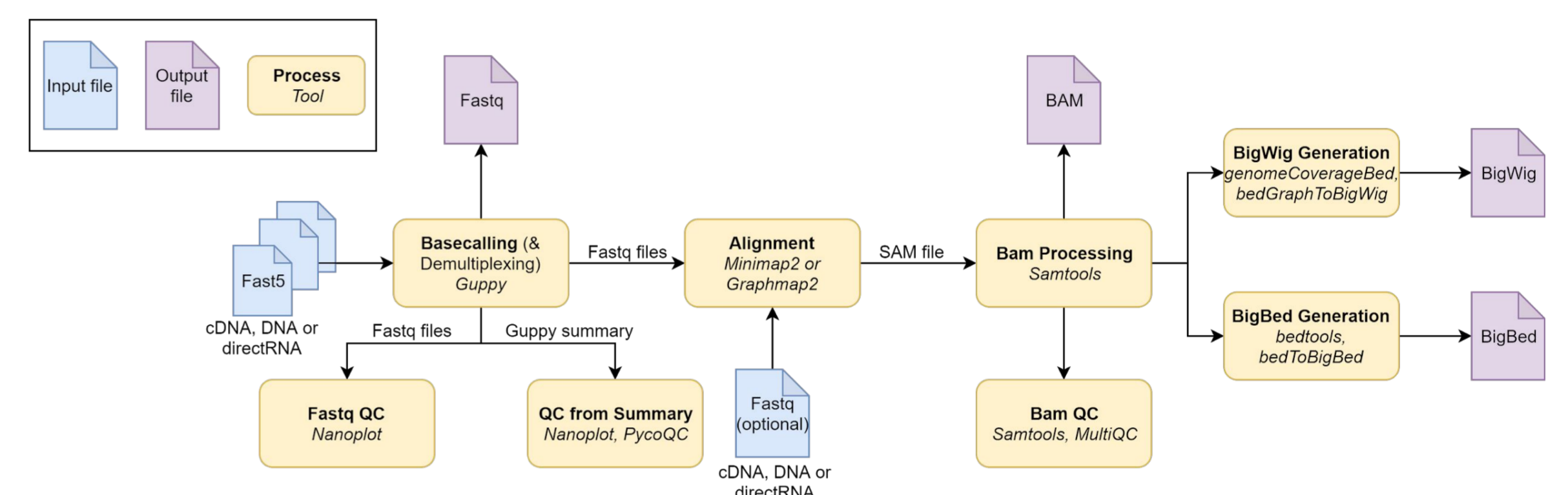
-> Human Transcriptomics Breakout Talk by Chen Ying



 [GitHub https://github.com/goekelab/sg-nex-data](https://github.com/goekelab/sg-nex-data)

Standardised Data Nanopore RNA-Seq Processing with Nextflow


nextflow nf-core



We have developed standardised pipelines in nextflow to simplify processing and analysis of Nanopore RNA-Seq data. The nanoseq pipeline enables processing and alignment of raw data. The nanorna pipeline facilitates downstream analysis (transcript discovery, quantification and differential expression analysis).

The nanoseq pipeline is developed by @csawye01, @drpatelh, @lwratten, @alneberg, @FranBonath and @AnnaSyme as part of the open source project nf-core.

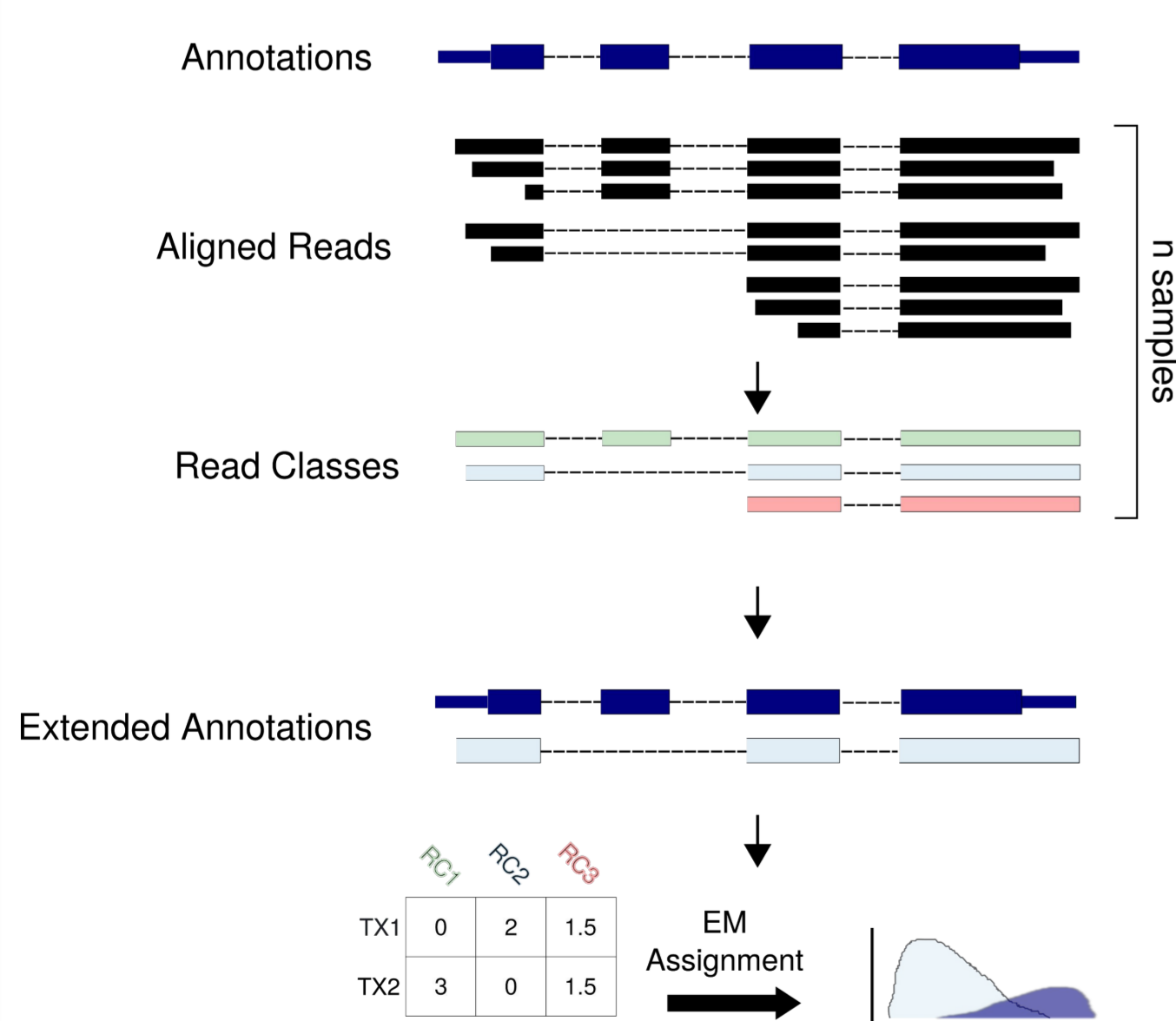
nf-core: Ewels PA, Peltzer A, Fillinger S, Alneberg JA, Patel H, Wilm A, Garcia MU, Di Tommaso P, Nahnsen S. nf-core: Community curated bioinformatics pipelines. *Nat Biotechnol.* 2020 Feb 13. doi: 10.1038/s41587-020-0439-x.

 [GitHub https://github.com/nf-core/nanoseq](https://github.com/nf-core/nanoseq)
<https://github.com/goekelab/nanorna-bam>

Transcript Discovery and Quantification

bambu

Developer: Chen Ying



- bambu is an R package for transcript discovery and quantification using Nanopore RNA-Seq data
- Allows multi-sample isoform discovery
- Performs quantification on annotated and extended transcript set
- Can be used directly for downstream analysis with other R packages such as DESeq2, DEXSeq

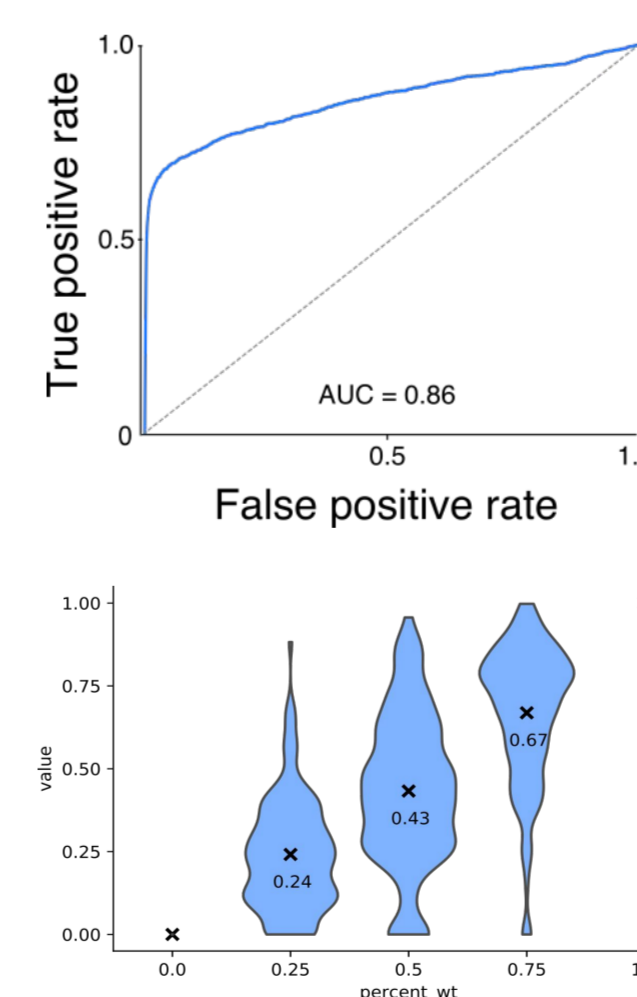


 [GitHub https://github.com/goekelab/bambu](https://github.com/goekelab/bambu)

Quantification of Differential RNA Modifications from Direct RNA-Seq

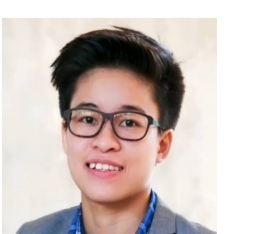
xPore

Developer: Ploy N Pratanwanich



- xPore is a method to detect differential RNA modifications from direct RNA-Sequencing data
- Raw signals are modelled by a multi-sample two-Gaussian mixture
- m6A sites are identified accurately
- Estimated modification rates correspond to fractions of modified RNAs

-> Data Analysis Theater Talk by Ploy N. Pratanwanich



 [GitHub https://github.com/goekelab/xpore](https://github.com/goekelab/xpore)