RNA-seq Bioinformatics:
Isoforms

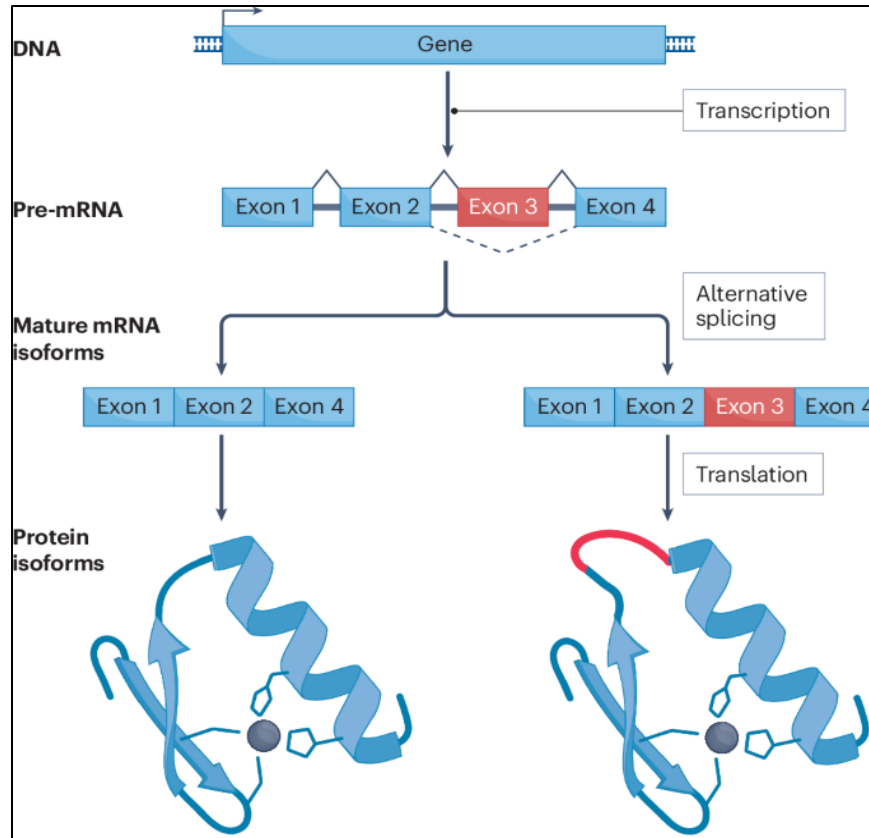Falko Noé

Isoforms

Hansen *et al*. 2024
PMID: 39232238

Isoforms
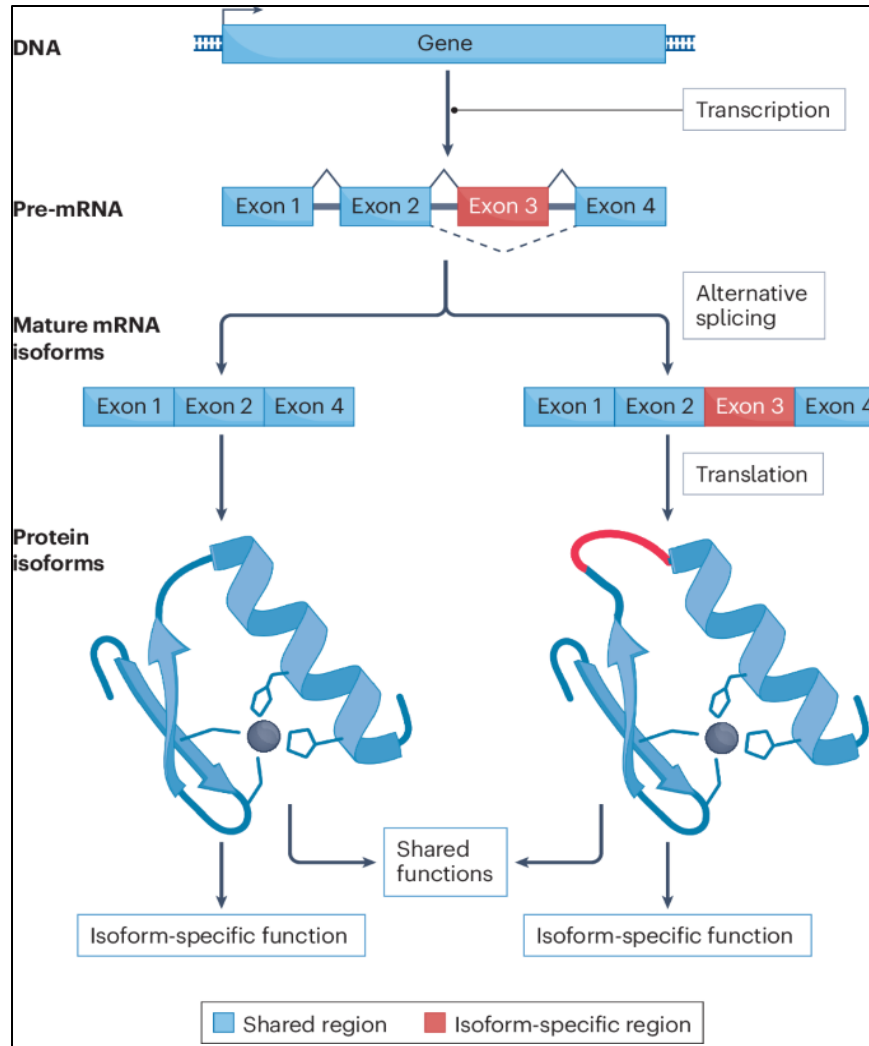
Hansen *et al.* 2024
PMID: 39232238

Transcriptomics in the era of long-read sequencing

Carolina Monzó [#] [1], Tianyuan Liu [#] [2], Ana Conesa [3]
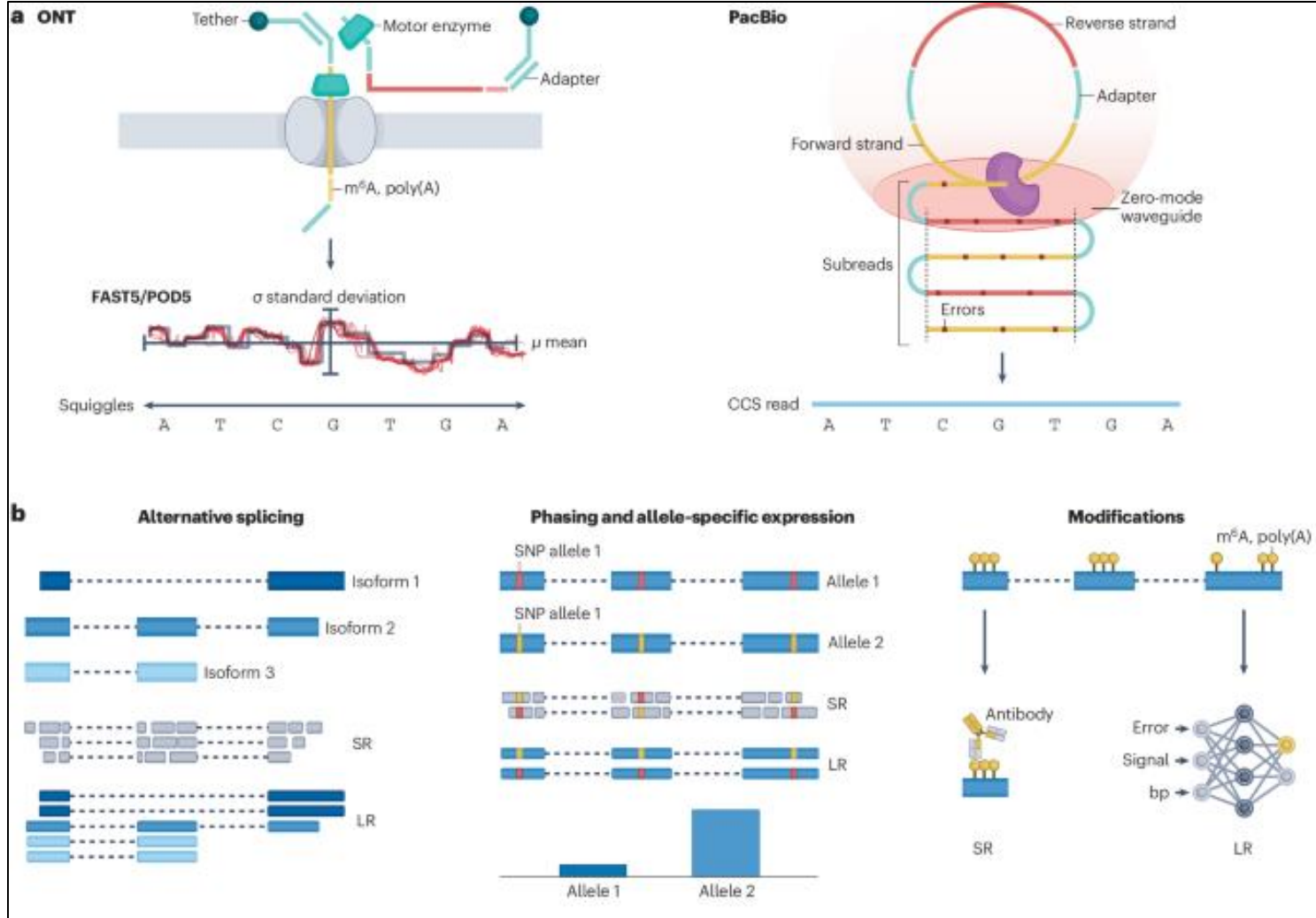
a ONT — Tether, Motor enzyme, Adapter, m⁶A, poly(A), FAST5/POD5, σ standard deviation, μ mean, Squiggles A T C G T G A

PacBio — Reverse strand, Adapter, Forward strand, Zero-mode waveguide, Subreads, Errors, CCS read A T C G T G A

b Alternative splicing — Isoform 1, Isoform 2, Isoform 3, SR, LR

Phasing and allele-specific expression — SNP allele 1, Allele 1, Allele 2, SR, LR, Allele 1, Allele 2

Modifications — m⁶A, poly(A), Antibody, Error, Signal, bp, SR, LR

Widespread variation in molecular interactions and regulatory properties among transcription factor isoforms

Luke Lambourne [1], Kaia Mattioli [2], Clarissa Santoso [3], Gloria Sheynkman [1], Sachi Inukai [4], Babita Kaundal [5], Anna Berenson [6], Kerstin Spirohn-Fitzgerald [1], Anukana Bhattacharjee [7], Elisabeth Rothman [4], Shaleen Shrestha [8], Florent Laval [9], Brent S Carroll [4], Stephen P Plassmeyer [10], Ryan J Emenecker [10], Zhipeng Yang [1], Deepa Bisht [5], Jared A Sewell [8], Guangyuan Li [7], Anisa Prasad [11], Sabrina Phanor [4], Ryan Lane [8], Devlin C Moyer [12], Toby Hunt [13], Dawit Balcha [1], Marinella Gebbia [14], Jean-Claude Twizere [15], Tong Hao [1], Alex S Holehouse [10], Adam Frankish [13], Josh A Riback [16], Nathan Salomonis [7], Michael A Calderwood [1], David E Hill [1], Nidhi Sahni [17], Marc Vidal [18], Martha L Bulyk [19], Juan I Fuxman Bass [20]
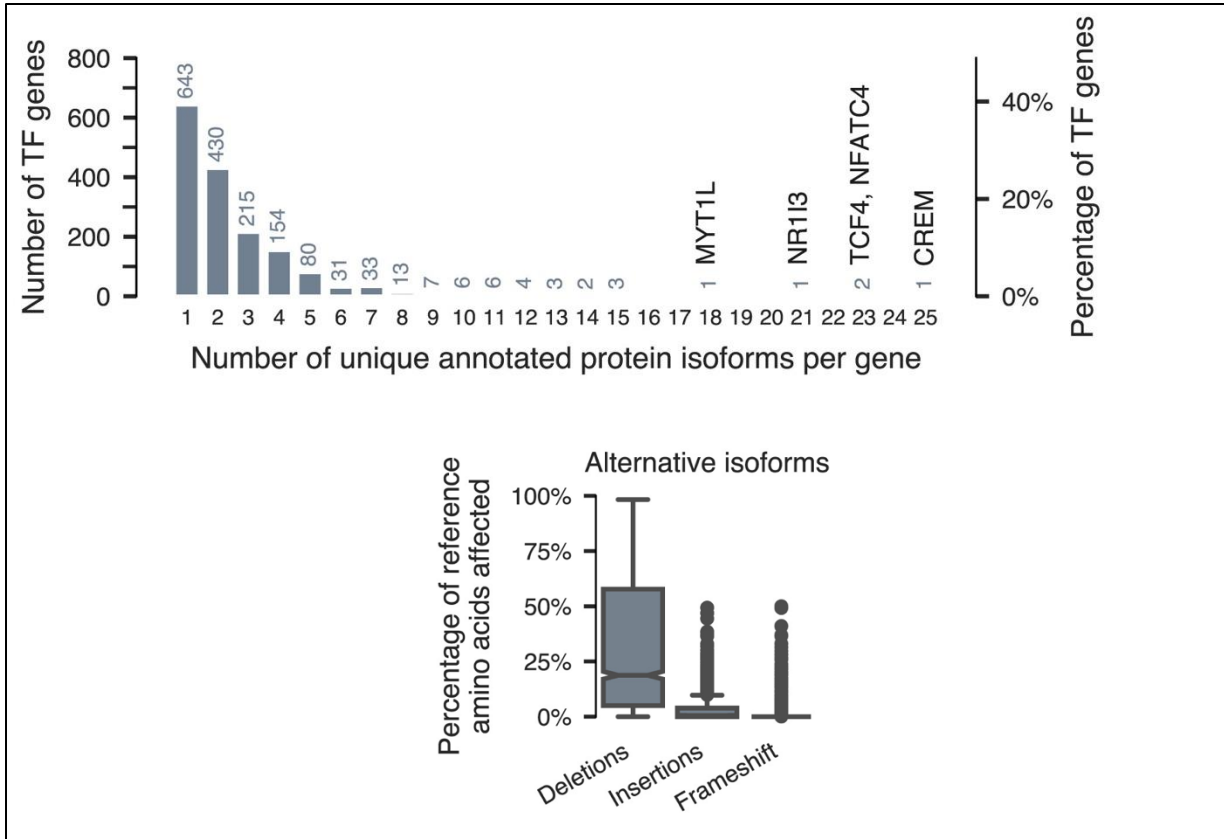
Lambourne et al. 2025
PMID: 40147441

Widespread variation in molecular interactions and regulatory properties among transcription factor isoforms

- Out of 246 TFs, hundreds of isoforms and 2/3 of them differ from their references in at least one key molecular activity, including DNA-binding or transcriptional effects

- ~50% of alternative TF isoforms show differential subcellular localization or condensate formation

Protein isoform-centric therapeutics: expanding targets and increasing specificity

Peter Kjer-Hansen [1] [2], Tri Giang Phan [3] [4], Robert J Weatheritt [5] [6]

- Expression of protein isoforms can be specific to a disease, tissue and/or stage

- Protein isoform-centric therapeutics can boost drug specificity

Long-read RNA sequencing atlas of human microglia isoforms elucidates disease-associated genetic regulation of splicing

Jack Humphrey [1,2,3,4,5], Erica Brophy [1,2,3,4,5], Roman Kosoy [1,2,4,6,7], Biao Zeng [1,2,4,6,7], Elena Coccia [1,2,3,4,5], Daniele Mattei [1,2,3,4,5], Ashvin Ravi [1,2,3,4,5], Tatsuhiko Naito [1,2,3,4,5], Anastasia G Efthymiou [1,2,3], Elisa Navarro [1,2,3,4,5,8,9,10], Claudia De Sanctis [2,3,11], Victoria Flores-Almazan [11], Benjamin Z Muller [1,2,3,4,5], Gijsje J L J Snijders [2,3,7], Amanda Allan [1,2,3,4,5], Alexandra Münch [1,2,3], Reta Birhanu Kitata [12], Steven P Kleopoulos [1,2,4,6,7], Stathis Argyriou [1,2,4,6,7], Periklis Malakates [1,2,4,6,7], Konstantina Psychogyiou [1,2,4,6,7], Zhiping Shao [1,2,4,6,7], Nancy Francoeur [1,4], Chia-Feng Tsai [12], Marina A Gritsenko [12], Matthew E Monroe [12], Vanessa L Paurus [12], Karl K Weitz [12], Tujin Shi [12], Robert Sebra [1,4,13,14], Tao Liu [12], Lot D de Witte [2,3,7], Alison M Goate [1,2,3,4,5], David A Bennett [15], Vahram Haroutunian [2,16], Gabriel E Hoffman [1,2,4,6,7], John F Fullard [1,2,4,6,7], Panos Roussos [#,17,18,19,20,21,22], Towfique Raj [#,23,24,25,26,27]

- Leverages long-read RNA sequencing to identify 35,879 novel microglia

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

- Contextual fear conditioning, a behavioral learning paradigm that induces differential gene expression in the hippocampus

- 1,266 significantly experience-variable isoforms identified

- Roughly 20% of the experience-variable isoforms arose from alternatively-spliced genes that showed no differential expression at the gene level

Transcript isoform diversity defines molecular subtypes and prognosis in acute myeloid leukemia through long-read sequencing

Xiaoguang Shi [1], Shuai Wang [2], Shuting Yu [2], Fangying Jiang [2], Jinzeng Wang [2], Yao Dai [2], Jincan Li [2], Ruihong Zhang [2], Tong Yin [2], Yuting Dai [2], Hai Fang [2], Junmin Li [2], Hongming Zhu [3], Shengyue Wang [4]

- Identified 119,278 previously unannotated transcript isoforms. Of these, 80,294 (67.31%) contain complete open reading frames, with 9,812 (12.22%) validated using MS

- Still, plenty of benchmarks and lack of consensus...

Systematic assessment of long-read RNA-seq methods for transcript identification and quantification

Francisco J Pardo-Palacios [#] [1], Dingjie Wang [#] [2] [3], Fairlie Reese [#] [4] [5], Mark Diekhans [#] [6], Sílvia Carbonell-Sala [#] [7], Brian Williams [#] [8], Jane E Loveland [#] [9], Maite De María [#] [10] [11], Matthew S Adams [12], Gabriela Balderrama-Gutierrez [4] [5], Amit K Behera [13], Jose M Gonzalez Martinez [9], Toby Hunt [9], Julien Lagarde [7] [14], Cindy E Liang [12], Haoran Li [2] [3], Marcus Jerryd Meade [15], David A Moraga Amador [16], Andrey D Prjibelski [17] [18], Inanc Birol [19], Hamed Bostan [20], Ashley M Brooks [20], Muhammed Hasan Çelik [4] [5], Ying Chen [21], Mei R M Du [22], Colette Felton [13], Jonathan Göke [21] [23], Saber Hafezqorani [19], Ralf Herwig [24], Hideya Kawaji [25], Joseph Lee [21], Jian-Liang Li [20], Matthias Lienhard [24], Alla Mikheenko [26], Dennis Mulligan [13], Ka Ming Nip [19], Mihaela Pertea [27] [28], Matthew E Ritchie [22] [29], Andre D Sim [21], Alison D Tang [13], Yuk Kei Wan [21] [30], Changqing Wang [22], Brandon Y Wong [27] [28], Chen Yang [19], If Barnes [9], Andrew E Berry [9], Salvador Capella-Gutierrez [31], Alyssa Cousineau [32], Namrita Dhillon [13], Jose M Fernandez-Gonzalez [31], Luis Ferrández-Peral [1], Natàlia Garcia-Reyero [33], Stefan Götz [34], Carles Hernández-Ferrer [31], Liudmyla Kondratova [35], Tianyuan Liu [36], Alessandra Martinez-Martin [1], Carlos Menor [34], Jorge Mestre-Tomás [1], Jonathan M Mudge [9], Nedka G Panayotova [16], Alejandro Paniagua [1], Dmitry Repchevsky [31], Xingjie Ren [37], Eric Rouchka [38], Brandon Saint-John [13], Enrique Sapena [39], Leon Sheynkman [15], Melissa Laird Smith [38], Marie-Marthe Suner [9], Hazuki Takahashi [40], Ingrid A Youngworth [41], Piero Carninci [40] [42], Nancy D Denslow [10] [43], Roderic Guigó [7] [44], Margaret E Hunter [45], Rene Maehr [32], Yin Shen [46], Hagen U Tilgner [47], Barbara J Wold [8], Christopher Vollmers [48], Adam Frankish [49], Kin Fai Au [50] [51], Gloria M Sheynkman [52] [53] [54], Ali Mortazavi [55] [56], Ana Conesa [57] [58], Angela N Brooks [59] [60]

AnaConesa @anaconesa · 7 jun. 2024 · · ·

We generated over 450 million long reads on 3 species with several library preparation and sequencing platforms, and drafted 3 challenges: transcript annotation with a reference annotation, transcript quantification, and transcript identification without a reference (2/6).

RGASP Challenge

AnaConesa @anaconesa · 7 jun. 2024 · · ·

14 tools submitted predictions that were analyzed by an independent team of evaluators applying different benchmarking strategies, including #SQANTI from @conesa_lab, @the_au_lab, @GencodeGenes curation, and @GSheynkman for experimental validation (3/6).

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

Long-read sequencing transcriptome quantification with lr-kallisto

Posted January 29, 2025.

Rebekah K. Loving, Delaney K. Sullivan, A. Sina Booeshagi, Fairlie Reese, Elisabeth Rebboah, Jasmine Sakr, Narges Rezaie, Heidi Y. Liang, Ghassan Filimban, Shimako Kawauchi, Conrad Oakes, Diane Trout, Brian A. Williams, Grant MacGregor, Barbara J. Wold, Ali Mortazavi, Lior Pachter

doi: https://doi.org/10.1101/2024.07.19.604364

Download PDF
Email
Print/Save Options
Share
Supplementary Material
Citation Tools
Data/Code
Get QR code
Revision Summary

Lior Pachter
@lpachter

···

We benchmarked some other tools that have recently been developed for long-read quantification, including Bambu, Oarfish, and IsoQuant. lr-kallisto is faster and more accurate. 6/

- Still, plenty of benchmarks and lack of consensus…

bioRχiv

THE PREPRINT SERVER FOR BIOLOGY

A Systematic Benchmark of High-Accuracy PacBio Long-Read RNA Sequencing for Transcript-Level Quantification

David Wissel, Madison M. Mehlferber, Khue M. Nguyen, Vasilii Pavelko, Elizabeth Tseng, Mark D. Robinson, Gloria M. Sheynkman

Posted June 30, 2025.

This article is a preprint and has not been certified by peer review [what does this mean?].

↗ **Download PDF**   ✉ Email
▽ **Print/Save Options**   ➤ Share
✎ Revision Summary   🌐 Citation Tools
                           ▦ Get QR code