

THE SQANTIVERSE TO EVALUATE AND CORRECT LONG-READ TRANSCRIPTOMICS DATA

About the speaker

Ana Conesa is a computational biologist and Research Professor at the Spanish National Research Council (CSIC), based at the Institute for Integrative Systems Biology in Valencia, and Courtesy Professor at the University of Florida. She is a member of the Spanish Royal Academy of Engineering, Fellow and Vice-President of the International Society for Computational Biology (ISCB), and President of the Spanish Society of Bioinformatics and Computational Biology. She directs the CSIC network for Computational Biology and the CSIC node of the European infrastructure ELIXIR, and co-founded Biobam Bioinformatics.

Her laboratory studies functional aspects of gene expression at the genome-wide level and has developed more than twenty statistical methods and software tools for transcriptomics, including Blast2GO, NOISeq, maSigPro, Qualimap, PaintOmics, tappAS and the SQANTI family, which together have received over 46,000 citations. A strong drive in her research is helping the genomics community bridge the gap between data and knowledge by building bioinformatics tools that everybody can use. She has led several international projects on multi-omics integration and on long-read transcriptomics, among them LRGASP, the community-wide benchmark of long-read methods for transcriptome analysis.



Prof. Ana Victoria Conesa Cegarra
Research Professor
Institute for Integrative Systems Biology (I2SysBio),
Spanish National Research Council (CSIC) -
University of Valencia, Spain



Friday, 18 September 2026
11am to 12pm (SGT, GMT+8)



GIS Seminar Room (Level 2)

About the seminar

Long-read sequencing promises to resolve transcriptomes at full-length isoform resolution, but long reads are noisy, and the transcript models we derive from them depend heavily on the library preparation, the sequencing platform and the reconstruction algorithm. Telling genuine biology from technical artifact has therefore become the central problem of long-read transcriptomics.

In this seminar, Ana will present the SQANTIverse, the ecosystem of methods her group has built around this problem. Starting from the structural classification and quality control of transcript models in SQANTI3, she will show how machine-learning curation yields high-confidence transcriptomes, how SQANTI-SIM simulates novelty for controlled benchmarking, and how TUSCO provides an internal control that reports on library preparation limitations and RNA degradation. Ana will then illustrate how the framework extends to single-cell data, genome annotation, protein isoforms, functional annotation and epigenomics, drawing on results from the LRGASP benchmarking effort.

Finally, Ana will argue that some of what we dismiss as noise is real biology, and introduce transcript divergence as a measure of unproductive mRNA biogenesis, which we find increases with age and under stress.