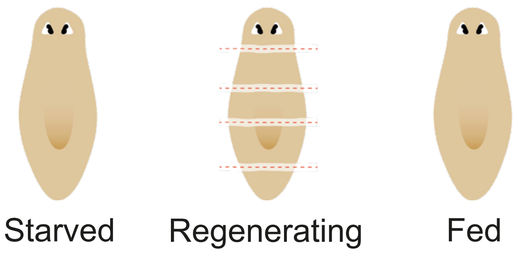
Long read nanopore RNA sequencing improves planarian transcriptome annotation and differential gene expression analysis approaches

Maria Rosselló [1,2,*], Teresa Adell [1,2], Emili Saló [1,2], Josep F Abril [1,2]

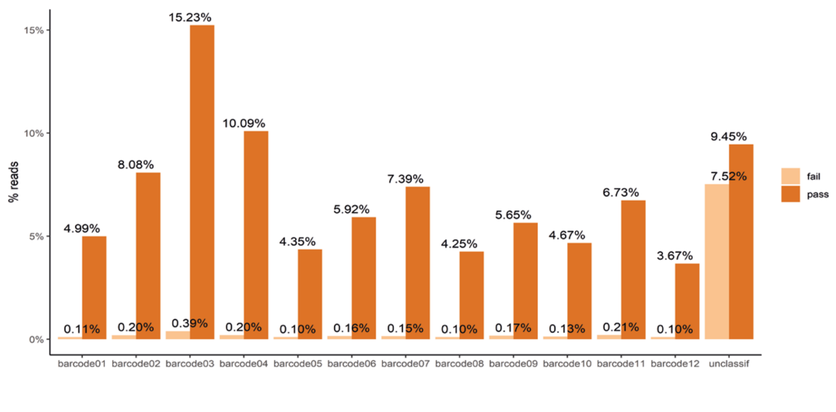
[1] Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia (UB), Avinguda diagonal 643, 08028, Barcelona, Catalonia, Spain.
[2] Computational Genomics Lab; Institut de Biomedicina (IBUB), Universitat de Barcelona, Barcelona, Catalonia, Spain.
[*] mariarossello@ub.edu

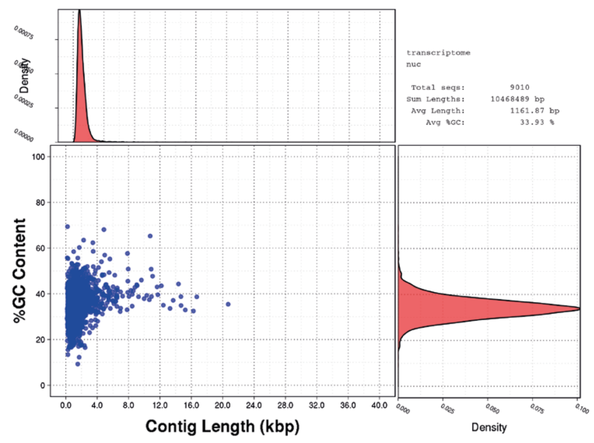
Abstract

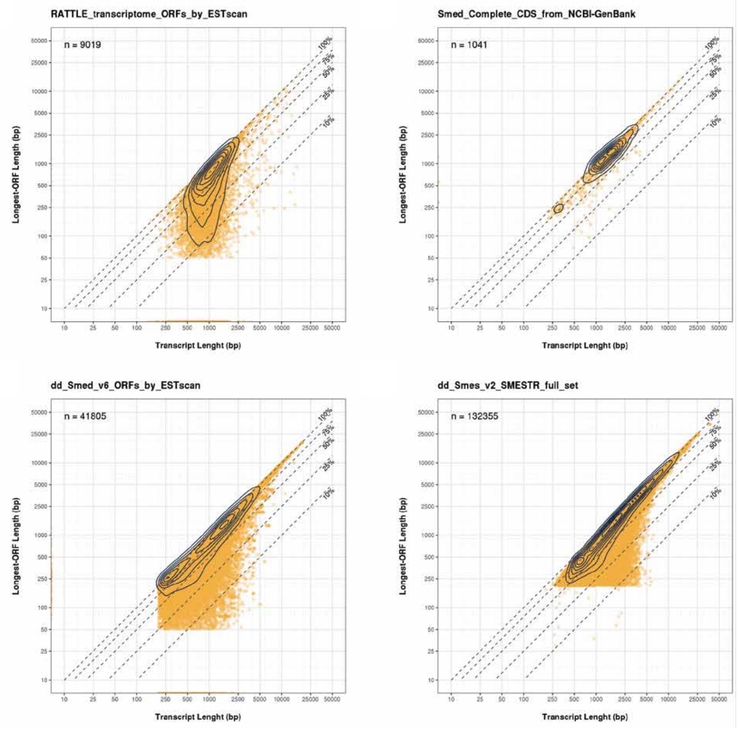
Planarians have become a model organisms in regeneration and adult tissue renewal research. Despite its relevance in that field, *Schmidtea mediterranea* genomic and transcriptomic diversity is not completely understood yet. In this work we propose a new approach to obtain full length transcripts using a single-molecule long-read sequencing method based on Oxford Nanopore technologies (ONT) for the first time in this species. Using nanopore sequencing data we could successfully assemble a transcriptome and describe some alternative splicing isoform events. We further implemented some improvements on the computational protocol to correct the errors derived from the sequencing methodology to retrieve the proper ORFs translations. From the multiplexed samples of our sequencing run, we demonstrated that data derived from nanopore sequencing can be applied to differential gene expression (DGE) analyses too. The results from the comparison among different physiological states in the planarian led us to undercover key genetic pathways, which play a role on the molecular mechanisms that control regeneration and body size.
In conclusion, in this work we established that single-molecule long-read sequencing can be a reliable method for de novo transcriptome assembly as well as DGE analyses in *S. mediterranea*.



de novo assembly

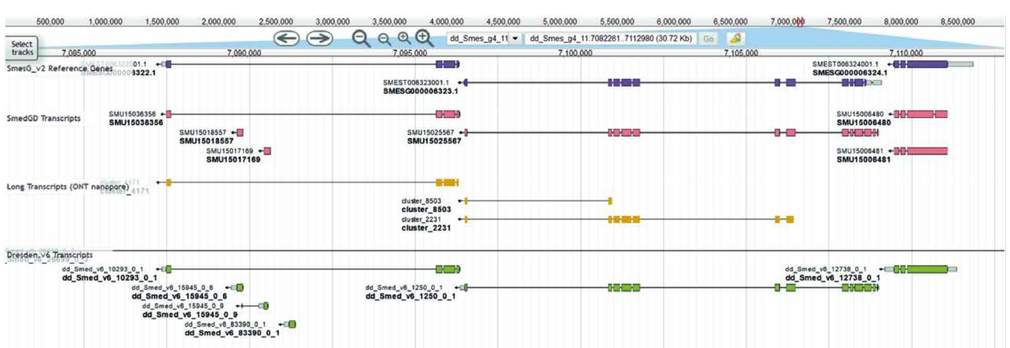


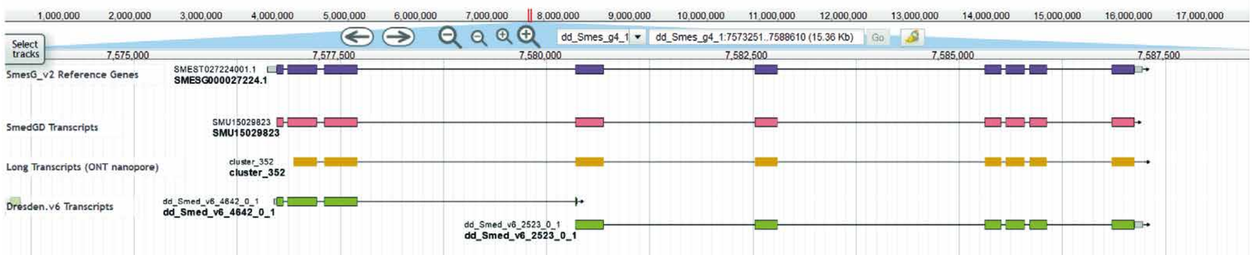




Nanopore sequencing enabled us to obtain contigs labeled with the barcodes assigned to each sample and the corresponding replicates. Using RATTLE we could assemble a *de novo* transcriptome for planarian that contains full length reads. We could successfully predict ORFs using ESTscan and the results were comparable with other transcriptomes previously assembled in planarians from Illumina RNA-seq experiments.

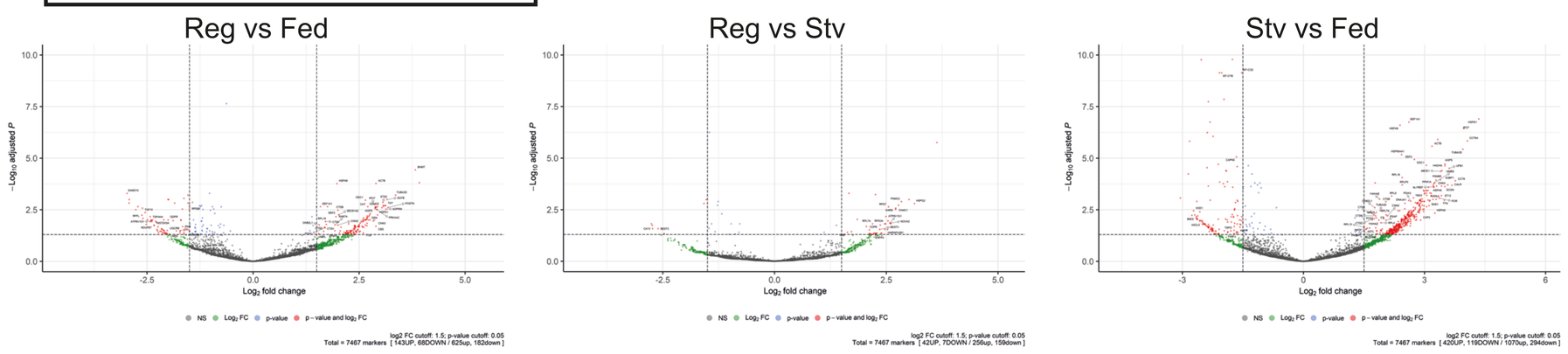
Transcript reannotation

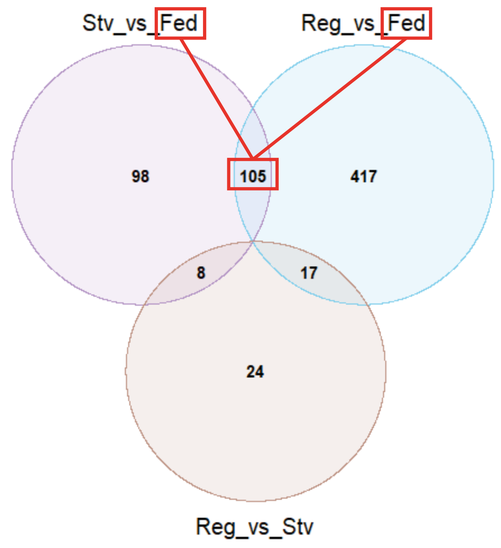




The newly obtained transcripts are useful to to explore new isoforms that can be represented in a JBrowse genome browser.

Differential gene expression





The data obtained using Nanopore sequencing was useful to perform a differential gene expression analysis (DGE). In this design , we compared different physiological conditions (starved, regenerating, and feeded) of the planarian and we could determine which genes are key in planarian regeneration and body remodeling.

BFU2017-83755-P