Universität Basel

Research Project

Cell type-specific expression of 3' untranslated region isoforms: quantification, modeling, and prediction of functional impact

Third-party funded project

Project title Cell type-specific expression of 3' untranslated region isoforms: quantification, modeling, and prediction of functional impact
Principal Investigator(s) Zavolan, Mihaela ;
Project Members Herrmann, Christina ; Katsantoni, Maria ; Bak, Maciej ; Schlusser, Niels ; Ghosh, Shreemoyee ; Moon, Youngbin ; Mironov, Aleksei ;
Organisation / Research unit
Departement Biozentrum / Bioinformatics (Zavolan)
Department
Project start 01.10.2019
Probable end 30.09.2023
Status Completed

The human body contains hundreds of cell types. Thus, a central challenge of modern biology is to uncover the principles that guide the formation and maintenance of this complex cellular landscape, from a single cell and based on the same genome sequence. Indeed, large scale initiatives such as the 'human cell atlas' are ongoing, aiming to comprehensively map cell states with a variety of technologies, including single cell sequencing. However, while most human genes express multiple isoforms, differing primarily in transcription start and termination sites, expression is mostly described at the 'gene level'. Furthermore, annotation of regulatory elements and their specific involvement in a particular cell remains sparse.

The term alternative cleavage and polyadenylation (APA) is used to describe the tissue-dependent choice of a poly(A) site among the many available within a gene. APA leads to expression of isoforms that differ in their protein-coding sequences (CDS) and/or in their 3' untranslated regions (3' UTRs). During animal evolution, 3' UTRs have expanded from 140 nucleotides in the worm Caenorhabditis elegans to 1-2 kilobases in humans, suggesting a parallel increase in the complexity of post-transcriptional regulation. Indeed, there is increasing evidence that APA contributes to cell identity. For instance, it has been reported that the cell type-specific pattern of expression of single-3'UTR genes is due preferentially to transcriptional regulation, and rather to changes in isoform ratios in the case of multi-UTR genes. Because 3' UTR isoforms differ in relative stability, localization, and translation rate, APA-dependent remodeling of 3' UTRs has many and far-reaching consequences on cell physiology. A very surprising finding has been that transcripts with distinct 3' UTRs can direct the localization of the (same) encoded protein to distinct cellular compartments. Although these findings made clear that much of the diversity of cellular phenotypes could come from differences in APA-dependent isoform expression, this possibility that has received relatively little consideration.

Having established unique resources and methods in the field of APA, I here propose to develop a coherent set of synergistic methods to systematically annotate cell type-specific poly(A) sites and isoforms, quantify their usage at the single cell level, and model the regulation of cell type-specific isoform expression. Because the broadly used 10x Genomics technology for single cell sequencing captures precisely transcript 3' ends, we will exploit such data in combination with comprehensive genome-wide annotations of alternative poly(A) sites to quantify APA in single cells. We will develop new methods to systematically infer regulatory motifs for a large set of RNA binding proteins (RBPs) using extensive eCLIP data from ENCODE. Using these RBP motifs in combination with RNA sequencing data from

The Cancer Genome Atlas, the human cell atlas and other consortia, we will model poly(A) site usage in terms of configurations of RBP binding sites, to infer how cell-type-dependent processing of poly(A) sites is regulated in physiological and pathological conditions. Finally, we will develop methods to assess the downstream effects of APA on RNA stability, localization, and translation. The resources generated in this project will be combined with data from large-scale initiatives, e.g. for personalized health, to link isoform expression to cellular morphology and phenotype, support identification of biomarkers, and the development of individualized therapies.

Keywords PWM, 3'UTR, modeling, single cell, Poly(A) site

Add publication

Published results

4626790, Burri, Dominik; Zavolan, Mihaela, Shortening of 3p UTRs in most cell types composing tumor tissues implicates alternative polyadenylation in protein metabolism, 1469-9001, RNA (New York, N.Y.), Publication: JournalArticle (Originalarbeit in einer wissenschaftlichen Zeitschrift)

4643034, Ghosh, Souvik; Ataman, Meric; Bak, Maciej; Börsch, Anastasiya; Schmidt, Alexander; Buczak, Katarzyna; Martin, Georges; Dimitriades, Beatrice; Herrmann, Christina J.; Kanitz, Alexander; Zavolan, Mihaela, CFIm-mediated alternative polyadenylation remodels cellular signaling and miRNA biogenesis, 0305-1048 ; 1362-4962, Nucleic Acids Research, Publication: JournalArticle (Originalarbeit in einer wissenschaftlichen Zeitschrift)

4665125, Katsantoni, Maria; van Nimwegen, Erik; Zavolan, Mihaela, Improved analysis of (e)CLIP data with RCRUNCH yields a compendium of RNA-binding protein binding sites and motifs, 1474-7596 ; 1474-760X, Genome Biology, Publication: JournalArticle (Originalarbeit in einer wissenschaftlichen Zeitschrift)

4697969, Moon, Youngbin; Burri, Dominik; Zavolan, Mihaela, Identification of experimentally-supported poly(A) sites in single-cell RNA-seq data with SCINPAS, 2631-9268, NAR genomics and bioinformatics, Publication: JournalArticle (Originalarbeit in einer wissenschaftlichen Zeitschrift)

Add documents

Specify cooperation partners