Methods xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

STaQTool: Spot tracking and quantification tool for monitoring splicing of single pre-mRNA molecules in living cells

José Rino *, Ana C. de Jesus, Maria Carmo-Fonseca

Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, 1649-028 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 2 November 2015
Received in revised form 2 February 2016
Accepted 4 February 2016
Available online xxxx

Keywords:
Live-cell imaging
Single-molecule
Fluorescence microscopy
Fluorescence quantification
RNA splicing

ABSTRACT

The vast majority of human protein-coding genes contain up to 90% of non-coding sequence in the form of introns that must be removed from the primary transcripts or pre-mRNAs. Diverse forms of mRNAs encoded from a single gene are created by the differential use of splice sites and alternative splicing is rapidly evolving. Although the kinetic properties of splicing are thought to be critical for proofreading and regulatory mechanisms, tools for making direct experimental measurements of splicing rates are still limited. We recently developed a strategy that permits real-time imaging of fluorescent-labelled introns in single pre-mRNA molecules. Here we describe the software tool that we created for automatic tracking and quantification of intronic fluorescence at the site of transcription in live human cells.

1. Introduction

A breakthrough in RNA imaging resulted from the development of a novel approach to visualize mRNA synthesis and transport in living mammalian cells that consists in genetically inserting binding sites for the MS2 bacteriophage coat protein in a gene of interest [1]. The resulting reporter transgene is then integrated in the genome of mammalian cells that express the MS2 coat protein fused to a fluorescent protein such as GFP. Upon transcription, binding of fluorescent coat protein to its target sequence makes that particular mRNA visible in living cells. Insertion of the MS2 binding sites in the terminal exon of reporter genes revealed kinetic properties of the entire mRNA life cycle, from transcription to transport in the nucleus and export to the cytoplasm [2,3]. More recently, insertion of binding sites for phage coat proteins in introns has been used to visualize splicing in real time [4,5]. These studies analyzed an ensemble population of pre-mRNAs synthesized from a gene cluster comprising multiple copies of the reporter gene. Thus, multiple nascent RNAs were simultaneously detected, necessitating a modelling approach to infer kinetic information. To circumvent these significant limitations and potential problems in data interpretation, we developed a strategy that permits direct tracking of single pre-mRNA molecules in live cells [6].

Here, we describe in detail the software tool that we created to track transcription sites in 3D and quantify their total fluorescence intensity over time in live human cells. In order to automate the multiple processing steps required for the analysis, we further developed an open-source user-friendly graphical user interface (GUI) that is available at https://imm.medicina.ulisboa.pt/en/servicos-e-recursos/technical-facilities/bioimaging/.

2. Overview of the method

STaQTool was developed for automated tracking of a transcription site in the cell nucleus, measuring its fluorescence intensity by Gaussian fitting and analyzing patterns of changes in fluorescence intensity over time (Fig. 1). STaQTool runs in any computer with either MATLAB or MATLAB Compiler Runtime and can be used by researchers without any background in programming. The software is a set of open-source MATLAB functions integrated into a single GUI (Fig. 2). The user-friendly GUI integrates all functions in a self-explanatory interface and allows for data visualization and easy adjustment of tracking and fitting parameters, as well as automated calibration procedures and analysis of either single transcription sites or multiple mRNA particles moving throughout the nucleus or the cytoplasm. Results such as coordinates of the transcription site and total fluorescence intensity (TFI) are saved to Excel files automatically and allow the user to pause and resume analysis at any time, as well as set up batch analysis of multiple files for unattended automated processing. STaQTool was developed to be used with TIFF and LOG files generated by the Slidebook

* Corresponding author at: Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, 1649-028 Lisboa, Portugal.

E-mail address: joserino@medicina.ulisboa.pt (J. Rino).

http://dx.doi.org/10.1016/j.ymeth.2016.02.005
1046-2023/

software from 3i (https://www.intelligent-imaging.com/slidebook.php) but the software can also generate the required LOG file (a text file with data acquisition parameters) with the necessary metadata and associate it with any standard TIFF file exported from other imaging systems.

After loading a time-lapse sequence, the first step is to track the diffraction-limited spot corresponding to the transcription site (Fig. 3). This is done interactively: the software can identify all diffraction limited spots in an image and the user clicks on the location of the relevant spot to define its coordinates in the first time point. The software can then be set to track it automatically. Importantly, whenever a spot is no longer detected (i.e., when previously synthesized transcripts were spliced or released from the site of transcription and no new pre-mRNA molecules were synthesized), the tracking algorithm forces the search area to remain in the location where the spot was last observed. If the data consists on a time-lapse sequence of z-stacks, tracking is performed in 3D.

Once tracking is completed, the TFI of the transcription site can be calculated (either at each time point or automatically for the whole time-lapse sequence) by performing a 2D Gaussian fit at the coordinates of the spot (Fig. 3). If the data is a sequence of z-stacks over time, then the Gaussian fit is performed at the Z plane corresponding to the highest intensity value. In addition to the TFI, the Gaussian fit also yields the Gaussian width W. The grid size for Gaussian fitting is automatically determined from the full width at half maximum (FWHM) of the point spread function (PSF), an experimental parameter specific for each microscopy setup that can be easily obtained by imaging sub-resolution fluorescent beads and open source software such as ImageJ (http://imagej.nih.gov/ij/).

Plots of TFI and W over time can be obtained for multiple time-lapse sequences in automated batch mode. Since multiple pre-mRNAs can be simultaneously present at the transcription site and imaged as a single diffraction-limited spot, an estimate of the TFI corresponding to a single transcript is required to determine the number of pre-mRNAs that are present at any given time point and to identify single pre-mRNA splicing events.

The next step in the protocol is thus to determine the TFI range for single transcripts. This can be achieved by imaging cells treated with a splicing inhibitor that causes release of unspliced pre-mRNAs from the site of transcription [6] or by imaging individual mRNAs tagged with the same number of MS2-binding sites at the 3′ UTR. Once released, individual mRNA molecules diffuse throughout the nucleus. As each molecule diffuses with a unique random trajectory, any diffusing spot corresponds to a single RNA. The software can automatically detect all diffraction-limited spots corresponding to single mRNAs in the nucleus and perform Gaussian fitting to determine the TFI and W values for each spot (Fig. 4). The distribution of TFI values measured for mRNA molecules diffusing in the nucleoplasm typically corresponds to a unique population and the range of TFI and W values that correspond to 68% (one standard deviation) and 95% (two standard deviations) can be automatically calculated with STaQTool (Fig. 5).

Following the calibration step, TFI values can be translated into average number of RNA molecules present at the site of transcription. Cycles of fluorescence increase and loss corresponding to the lifetime of a fluorescently labelled intron in a single pre-mRNA molecule can then be automatically detected. STaQTool performs automatic detection of such events by searching for specific fluo-

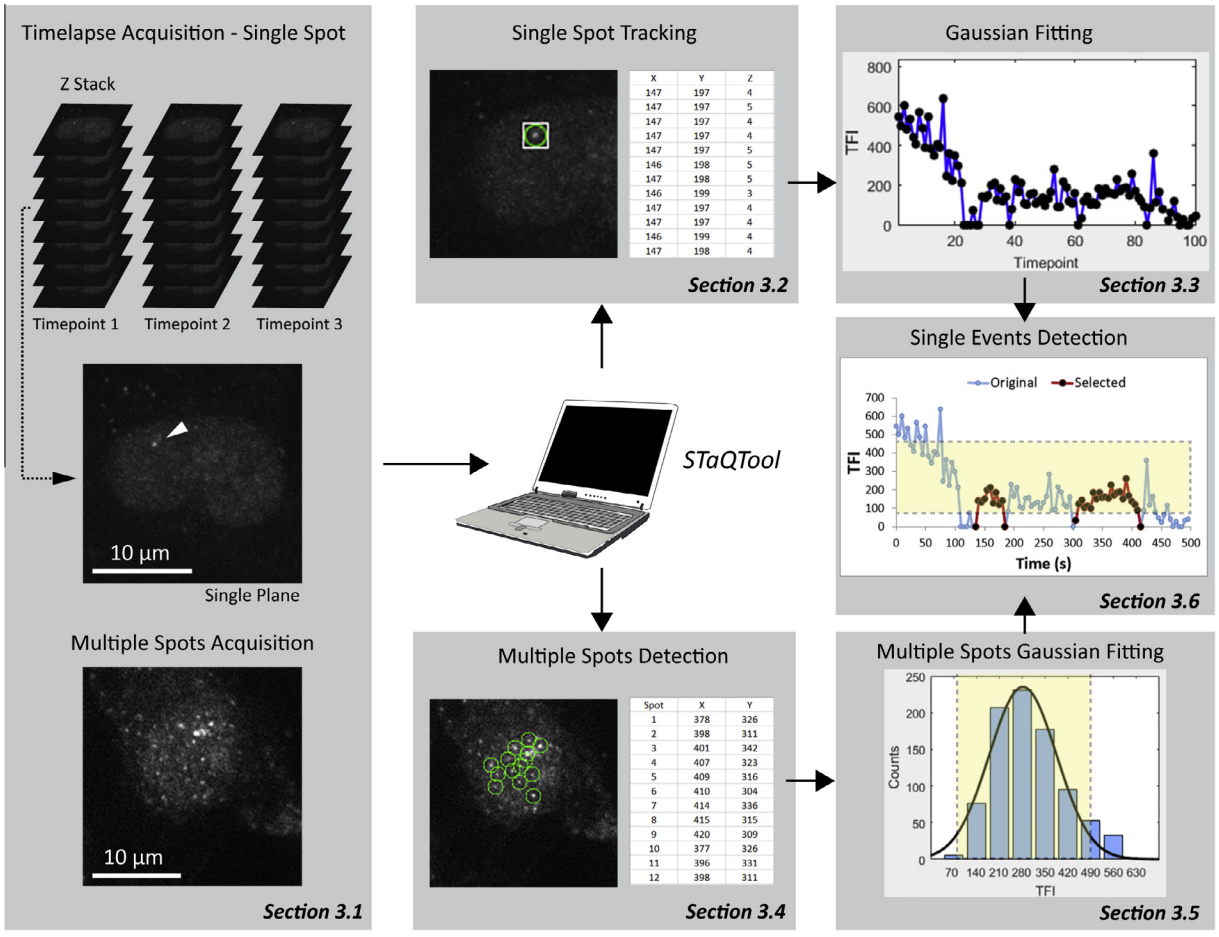

Fig. 1. STaQTool's workflow. The major processing steps are schematically illustrated with reference to the corresponding Software section.