Identification of SNA-I-positive cells as stem-like cells in an established cell line using computerized single-cell lineage tracking

Sachiko Sato[1], Ann Rancourt[1,2], and Masahiko S. Satoh[2]*

[1]Glycobiology and Bioimaging Laboratory of Research Center for Infectious Diseases, and

[2]Laboratory of DNA Damage Responses and Bioimaging, Research Centre of CHUQ,

Faculty of Medicine, Laval University, 2705 Boulevard Laurier, Quebec, G1V 4G2,

Canada

*Correspondence to Masahiko S. Satoh

TEL: 1-418-525-4444 ext. 47340

FAX: 1-418-654-2159

e-mail: Masahiko.sato@crchul.ulaval.ca

Abstract

Single-cell tracking analysis is a potential research technique for the accurate investigation of cellular behaviors and events occurring within a cell population. However, this analysis is challenging because of a lack of microscope hardware and software suitable for single-cell tracking analysis of a wide range of cell types and densities. We therefore developed a computerized single-cell lineage tracking analysis system based on a microscope optimized for differential interference contrast-based long-term live cell imaging, with software designed to automatically generate live cell videos, perform image segmentation, carry out single-cell tracking, and create and analyze a cell lineage database. We previously reported that minor cell sub-populations (3%–7%) within a cultured cancer cell line could play a critical role in maintaining the cell population. Given that sub-population characterization requires large-scale single-cell tracking analysis, we tracked single cells using the above computerized system and identified a minor cell population (1.5%) composed of *Sambucus nigra* agglutinin-I-positive cells, which acted as stem-like cells for the established culture. These results demonstrate the potential value of this computerized single-cell lineage tracking analysis system as a routine tool in cell biology, opening new avenues for research aimed at identifying previously unknown characteristics of individual cultured cells with high accuracy.

Introduction

Spatiotemporal data created by single-cell lineage tracking analysis of individual live cells contain information that defines the fate of individual cells and the characteristics of the cell population. These data allow the detection of rare cellular events occurring within a cell population, and the precise determination of the spatiotemporal relationships of progeny cells with their siblings and progenitors (*1*). Given that any cell population may exhibit cell heterogeneity, events occurring in certain individual cells or sub-populations of cells can be analyzed accurately using spatiotemporal data (*1*). In contrast, most cell biological studies have used population-based end-point assays that analyze average responses of cells at one or several specific time points, and the cellular events that occurred in those cells are then deduced using the average end-point data to reveal the process of cell differentiation or the responses of the cells to test substances. However, these analyses cannot provide accurate conclusions if a substance or treatment induces non-stoichiometric cellular responses and/or affects only a handful of cells due to the existence of cellular heterogeneity. Analytical systems that can extract unbiased spatiotemporal data for individual cells, and approaches able to characterize cells at a single-cell level, are therefore required to ensure the accurate investigation of cellular events over time, the process of cell differentiation, and cellular responses to various external substances, including carcinogens, mutagens, drugs, infectious microorganisms, and growth stimulants.

Tracking of individual cells and subsequent analysis of the results are tedious and involve the creation of long-term live cell videos, identifying and tracking individual cells in the videos, determining cellular events occurring in the individual cells, and analyzing the data. This type of analysis was previously performed by creating a live cell movie using 16 mm film (*2*), though the introduction of computerized live cell imaging, accompanied by software able to track individual cells, has now begun to generate spatiotemporal data for individual cells (*3-6*). However, single-cell tracking using a fluorescence-based live cell imaging approach can be hindered by phototoxicity, particularly when long-term cell tracking is required (*7*). Although phase contrast- and differential interference

contrast (DIC)-based imaging are versatile methods and are less cytotoxic compared with fluorescence-based imaging, tracking cells using phase contrast and DIC images remains challenging, since the image quality, cell type, and cell density can significantly influence the cell tracking accuracy (*5*). Consequently, systems that fully characterize individual cells by creating spatiotemporal data have not yet been well developed.

These difficulties could be overcome by developing a system that creates optimal DIC images for digital partitioning using a segmentation approach and for single-cell tracking, and performs image processing and data analysis using software optimized for single-cell tracking analysis. Here, we report the development of a computerized single-cell lineage tracking analysis system using DIC imaging (Fig. 1), and the results of analyses of individual HeLa cervical cancer cells using this system. The computerized single-cell lineage tracking analysis system is controlled with two computers: one dedicated to controlling the microscope, ensuring system stability during image acquisition, and another that receives all the imaging data, creates live cell videos, performs DIC image segmentation, tracks individual cells, and analyzes the data. The system also builds a live cell imaging video archive, object segmentation library, and cell lineage database using a series of C/C++/Objective-C-based software modules (Supplementary Table 1). These data resources allow various analyses, from the individual cell to population level, by linking video images and cell tracking data. We used this system to investigate the role of a minor sub-population in the maintenance of the HeLa cell population (*1*). Our results suggest that a sub-population comprising *Sambucus nigra* agglutinin-I (SNA-I) positive cells (SNA-I^+) possesses stem cell-like characteristics. It has been suggested that pancreatic cancer stem cells or inducible pluripotent stem cells, but not differentiated cells, are SNA-I^+ cells (*8-11*). We therefore propose that the HeLa cell line provides an *in vitro* model for the growth of stem-like cells, which play a critical role in the maintenance of cell population homeostasis in HeLa cells.

Results

A microscope for long-term live cell imaging.

The stability of the microscope and controlling computer are critical for successful imaging of cells during routine long-term (>1 week) live cell imaging and simultaneous monitoring under different culture conditions. We previously performed DIC-based long-term live cell imaging for 1–2 weeks using a commercially available microscope (*1*), with a failure rate of >80%. The causes of failure included reduced image quality during imaging, computer system crashes, and imaging software shutdown. We therefore designed a microscope system optimized for long-term live cell imaging by selecting the appropriate apparatus and modifying an existing microscope. Near-infrared (NIR) DIC (Supplementary Fig. 1) was used to obtain uniform image quality during imaging and to minimize phototoxicity, and an image-acquisition computer (Computer 1) was linked to a second computer (Computer 2) (Fig. 1) to stabilize Computer 1 and to allow Computer 2 to perform real-time image processing after transferring image files created on Computer 1 using *File Transfer* software (see Materials and Methods, "Microscope", for detail). We performed more than 200 long-term live cell imaging runs using the microscope with this data transfer system, with no imaging termination due to computer crashes, suggesting that long-term live cell videos could be stably created using this system.

Generation of optimal DIC images for image segmentation and single-cell tracking.

Image quality is a critical factor for accurate single-cell tracking (*5*). However, image quality varied between acquisitions and changed over time due to various causes, e.g. changing cell density and the emergence of cell debris. It was thus often difficult for the microscope to continuously create ideal images for single-cell tracking and image segmentation. We therefore developed a series of software programs (Supplementary Table 1) to generate live cell videos with optimal images for segmentation and single-cell tracking (see Materials and Methods, "Starting image acquisition", for details of image processing). *Image Processing Controller* software controlled the process of generating live cell videos

by coordinating the performance of the other software programs (Fig. 1). *Map* software allowed the x and y positions and the z-focused optical plane to be set for each field of view (FOV) and a two-dimensional (2D) image acquisition array (Supplementary Fig. 1b and 2a). Once the multi-layer TIFF image files were created by the microscope and saved on Computer 1, *File Transfer* immediately transferred the files to Computer 2. *Name Assignment* software then assigned a unique file name to each file to initiate image processing. Focused images were then created using *Focal Image Selection* software (Supplementary Fig. 2b-d), and *Contrast Set* software generated a customized background pattern, adjusted the background, stitched the images, and built a live cell imaging video archive (Supplementary Fig. 3), which was backed-up using *Data Backup* software (Fig. 1). Notably, stitched images were prepared for segmentation by filling blank parts without cells by setting each pixel value to 100 on a grayscale of 0–255, with values of 95–99 and 101–105 converted to 94 and 106, respectively. We tested this approach and confirmed that live cell imaging videos with normalized background were created automatically (Supplementary videos 1 and 2).

Segmentation of NIR-DIC images.

We developed *Outline Drawing* software to perform image segmentation. It was challenging to segment the NIR-DIC images (grayscale images), particularly when the image contained areas of high cell density that prevented it from being partitioned into regions. To overcome this problem, we developed a method referred to as Stepwise Area Expansion (SAE), which used the original NIR-DIC image and four images (TH1–4) created by applying four different thresholds (high to low) to the NIR-DIC image. Figure 2a shows some typical example images for low- and high-density cells, and the subsequent segmentations. First, pixels above the threshold value (e.g. 200, 170, 140, and 110 of 0–255 grayscale for TH1, 2, 3, and 4, respectively) were extracted from the NIR-DIC image and the extracted pixel values were converted to 150. Segmentation was then performed step by step, starting from the TH1 image (Fig. 2b). Connectivity analysis was performed on the TH1 image to identify groups of

pixels that were attached together, and an edge circle that surrounded the group of pixels was then determined (Fig. 2b-A, pink circles). The TH1 edge circles were overlaid on the corresponding NIR-DIC image (Fig. 2b-B). The pixels composing the circles were referred to as the original pixels. Using the overlaid NIR-DIC image as a reference, we examined the pixel values for each original pixel. The values of pixels located in a circle from 12 o'clock from the original pixels were examined (Fig. 2b-B and Line extension (magnified)). The edges of the pixels were extended until the values became either 100 (representing the NIR-DIC background), or 50 above or below the original pixel, as long as those extensions were towards the outside of the edge circles. The final pixel positions were determined by this technique. This extension was continued for 20 pixels distant from the original pixel. Similar examinations were performed for the other three directions (Fig. 2b-B: 3, 6, and 9 o'clock directions; yellow, light blue, and blue lines, respectively) and the final pixel positions were then linked to create four directional lines corresponding to the 12, 3, 6, and 9 o'clock directions (Fig. 2b-C and Edge linking (magnified), L 12', L3', L6', and L9'). Finally, the four lines were linked to make an edge circle (Fig. 2b-D, green circles), which completed the initial segmentation for the first TH1 image.

The TH1 edge circles were then overlaid on the TH2 image and another connectivity analysis was performed, except for the TH1 edge circle areas (Fig. 2b-E, green circles). Two types of edge circles (Fig. 2b-E, green and pink circles) were thus created at this stage. The pink circles that emerged in the TH2 image were overlaid on the corresponding NIR-DIC image, the pixel value of each original pixel in the pink circles was examined, the final pixel positions were determined (Fig. 2b-F), and the positions were linked to create four lines (Fig. 2b-G) and a new edge circle was created (Fig. 2b-H), as described for TH1. We applied a different expansion approach for the green edge circles that had been determined during TH1 processing (Fig. 2b-E, green circles): the location of each pixel on the edge of the circles (Fig. 2b-K) was moved by 1–4 pixels outside the original (Fig. 2b-L, Edge expansion (magnified)) to create the red circles (Fig. 2b-L). Both the green TH2 (Fig. 2b-H) and red TH2 circles (Fig. 2b-L) were then overlaid onto the TH2 image (Fig. 2b-I) and circle overlaps were removed to

create new edge circles (Fig. 2b-J). Those new edge circles were then applied to TH3, and the above-mentioned processes were repeated for TH3 and TH4 images. Areas surrounded by edge circles were numbered, as shown in Fig. 2c–f. It should be noted that, in some cases, one cell contained multiple areas and at least one area was required to be assigned to perform tracking of a target cell. For myotubes (Fig. 2f), multiple areas were assigned to one myotube, while in the case of a confluent cell culture, one area tended to be assigned to one cell (Figs. 2d and e). Because the level of segmentation can be adjusted by setting the minimal area size and values of TH1–4, it was also possible to assign a larger segmentation area to a cell (Fig. 2c). These results suggest that NIR-DIC images with various cell densities can be segmented using SAE. Finally, an object segmentation library was created from the segmented data.

Controlling object tracking.

We developed *Object Tracking Controller* software to control the processes of tracking segmented areas representing individual cells, interpreting the single-cell tracking data, and creating the cell lineage database (Fig. 3a). Videos stored in the live cell imaging archive and images at time point 1 or at the time point where segmented area tracking was started were used to assign cell lineage numbers to the segmented areas (Fig. 3b, green edge circles, which surround a segmented area) recorded in the object segmentation library. To temporally assign the number, *Object Tracking Controller* evaluated each segmented area to determine whether the area was either a possible tracking target, representing a cell, or a non-target, based on area size and brightness (Fig. 3b). Edge circles, indicated by green and yellow arrows, were possible targets and non-targets, respectively (Fig. 3b, Tracking target and non-target). Alternatively, an area representing a cell could be determined interactively by selecting and reshaping an area using the line and circle drawing tools (Fig. 3b, Tracking target (cell), red arrow). Both a cell lineage number (Fig. 3b, Lineage no. assignment, orange numbers) and cell number 0 were then assigned to an area representing a possible target or to a cell. Before starting the automated

segmented area tracking, *Object Tracking Controller* was used to verify if the choice of segment, i.e. the possible target, was appropriate for segmented area tracking. Verification was performed by comparing segment area patterns for 10 time points during automatic object tracking. Based on this analysis, *Object Tracking Controller* determined if a segmented area needed to be removed or changed to another segmented area that better represented a cell.

Once all segmented areas representing each cell were annotated, the cell lineage numbers and cell numbers were entered into a *Queue* list using *Object Tracking Controller* (Supplementary Fig. 4a), and automatic tracking of the segmented area was then performed. We commenced the tracking with 100–500 segmented areas at time point 1, resulting in tracking of 2,000–10,000 segmented areas due to cell proliferation (4–5 cell doublings) during the observation period. Each segmented area was tracked for 10–20 consecutive time points in images from the segment recorded at the top of the *Queue* list. Once the last segment recorded in the *Queue* list was processed, the automatic tracking process was repeated from the segment recorded at the top of the *Queue* list until 100–200 time points had been recorded (Fig. 3a). The automatic tracking processing speed was approximately 0.3 s/time point/segmented area (MacBook Pro, 2016, OS 10.12.6, 2.9 GHz Intel Core i7, solid state drive), and automatic tracking of 420 segmented areas representing progenitor cells and segmented areas corresponding to their progeny (11,640 cells) up to time point 693 (6930 min) took around 150 h. When automatic tracking for a given segmented area failed, the information for that area was entered into a *Check* list instead of the *Queue* list (Supplementary Fig. 4a). After processing 100–200 time points, the tracking data entered in the *Queue* and *Check* lists were evaluated to determine if these data accurately represented a cell, using the interactive functions of *Object Tracking Controller*. If necessary, the areas were readjusted using the interactive line and circle drawing tools (Supplementary Fig. 4b). Because automatic segmented area tracking accuracy never reaches 100% in practice, the interactive functions were essential to relate the tracking data for segmented areas to individual cells to build an accurate cell lineage database, which contains data regarding time points, cell lineage numbers, cell numbers, x and y positions of the cells,

types of events that occurred in the cells, sizes, average brightness, and gravity centers of the areas, information on the edge circles, fluorescence intensities of the cells, types of fluorophore, relationships of parent cells and their progeny, and area numbers assigned by *Outline Drawing*. While tracking 1,000 cells was sufficient for accurate statistical analysis (*1*), we performed single-cell tracking of 11,640 cells, which required about 100 h of interactive evaluation. *Object Tracking Controller* comprising automatic tracking and interactive functions-assisted evaluation thus demonstrated the capacity for routine single-cell lineage tracking analysis in the laboratory.

Automated object tracking.

We previously performed single-cell tracking manually (*1*) by focusing on a cell at a given time point and then looking for a cell with a similar shape in the image at the next time point, by evaluating the cell and its surrounding area. *Automatic Object Tracking* software mimics this process of comparison carried out by the human brain to track one segmented area at a time. To recognize cells, *Automatic Object Tracking* required guidance to determine if the digital signals in live cell imaging videos were related to the cells. We therefore used segmented area data stored in the object segmentation library to track individual cells. Figure 3c summarizes the tracking process of the segmented area carried out by *Automatic Cell Tracking*. Orange characters/numbers indicate segmentations assigned to cells in the time A image, and black numbers in the time A+1 image indicate unassigned segments (Fig. 3c-A and D, respectively). To track a segmented area representing a cell, the unassigned segments were reorganized based on information for the target and surrounding segmented areas from time A by dividing or merging these segments found in time A+1. For example, when the light orange cell in Fig. 3c-B (arrow) was tracked, the segment inside the cell (red circle and red arrow) was used to represent the cell (Fig. 3c-C). This red segment disappeared at time A+1 (Fig. 3c-E), and new segments (Fig. 3c-E, yellow circles 1 and 2) appeared. Yellow circle 2 partly overlapped with the target cell (Fig. 3c-E). Thus, to track the target cell, the segment representing the cell at time A+1 was

redefined as follows. First, the red segment for time A was overlaid on the time A+1 image (Fig. 3c-F) and the red segment was overlapped with the yellow segment (yellow circle 2) (Fig. 3c-F). Second, to guide the readjustment of the target segmentation area, surrounding information for the target cell, such as other neighboring segments (pink circles 1 and 2) at time A (Fig. 3c-C), were overlaid onto the time A+1 image (Fig. 3c-G). Because yellow circle 1 was not attached to the red circle and only overlapped with pink circle 1, yellow segment 1 was not considered to be part of the target cell. In contrast, since yellow circle 2 overlapped with both pink 2 and the red segment (Fig. 3c-G), the segment was divided in two (white circles 1 and 2, Fig. 3c-H). The redefined time A+1 segmentation was drawn by combining the red segment at time A with white segment 2 (Fig. 3c-I) as a segment representing the cell. Another example of the reorganization was performed by merging segments. Figure 3c-J shows that the ivory line-marked cell was represented by one segment (Fig. 3c-J, light blue circle). At time A+1, the area was divided into several small segments (Fig. 3c-K). After applying the light blue segment onto the time A+1 image, the small segments were merged to form a single segment (Fig. 3c-L, blue circle) representing the cell. These processes were repeated automatically for all cells in the *Queue* list.

Object Tracking Controller controlled these processes by sending information for a cell in the *Queue* list to *Automatic Object Tracking*. After tracking the cell for 10–20 time points (Supplementary Video 3) or when an error occurred during tracking, *Automatic Object Tracking* returned the tracking information to *Object Tracking Controller*. *Automatic Object Tracking* also detected bipolar, tripolar, and tetrapolar cell divisions, and cell death, cell fusion, and mitosis (Supplementary Fig. 5, and see Materials and Methods, "Automatic and interactive setting of cellular events", for detail), and the detected events or event-detection errors were returned to *Object Tracking Controller*. The automatic tracking error was 1.3 errors per 100 time points per cell, but the error rate varied depending on the cell type and density. Almost 98.7% of segmented areas that represented cells in the recorded video were tracked and the tracking data were successfully interpreted into single-cell tracking data, suggesting

that this automatic object tracking approach could be used for tracking individual cells. By combining the interactive functions of *Object Tracking Controller* to correct errors, this computerized system allowed the transfer of various layers of information from time-lapse movies into digitized data, which could then be used to create accurate cell lineage databases and analyze parameters, such as cell proliferation rate, doubling time of individual cells, cell motility, and the frequency of aberrant cellular events such as cell death, fusion, and multipolar cellular divisions. We characterized SNA-I$^+$ cells using the computerized single-cell lineage tracking analysis system and analyzed the data using *Data Analysis* software (Supplementary Fig. 5, and see Materials and Methods, "Data analysis software", for detail).

Presence of SNA-I$^+$ cells within HeLa cell populations.

We previously reported that HeLa cell populations contain fast-growing cells (*1*). Because this sub-population plays a role in maintaining the overall cell population, we tentatively referred to these fast-growing cells as putative cancer stem cells (*1*), though whether the cells have stem-cell like functions remains to be determined. Most HeLa cells express the cancer marker TAG-72 (sialyl Tn antigen) (*12-14*), the stem cell marker CD133 (*15-17*), and the stem cell glycan marker fucose α1-2galactose β1-3-structure, which is recognized by the rBC2LCN bacterial lectin (*18*) (Fig. 4a). Interestingly, SNA-I, which binds to oligosaccharides containing the sialic acid α-2-6 galactose structure (α2-6Sia) (*10, 11, 19-22*), recognized a group of HeLa cells within the population (Fig. 4a). Given that the α2-6Sia epitope detected by SNA-I is expressed in some types of cancer stem cells (*10*) and pluripotent stem cells (*20*), we investigated if SNA-I$^+$ cells were associated with putative cancer stem cells in HeLa cell populations.

Single-cell tracking of HeLa cells.

Based on the numbers of SNA-I$^+$ cells, we estimated that SNA-I$^+$ cells constituted only 1%–2% of the total HeLa cell population. We therefore employed the computerized single-cell lineage tracking analysis system to track a large number of HeLa cells, which were subjected to live cell imaging for 115.5 h (693 time points). The cover glass Lab-TekII 8-well chamber was then removed from the microscope stage, the cells were fixed and stained with tetramethylrhodamine (TRITC)-labelled SNA-I and DAPI, and the chamber was replaced on the stage for fluorescent imaging. The fluorescent images were incorporated into the live cell imaging video archive (Supplementary Fig. 6, and see Materials and Methods, "Staining of cells with SNA-I-TRITC and adjustment of image position after fluorescence imaging", for detail). The analysis created 420 cell lineage data by tracking 11,640 cells. The images and results of single-cell lineage tracking at times 1 and 639, as well as SNA-I$^+$ cells found in the HeLa cell population at time 639, are shown in Supplementary Fig. 7a, b, and c, respectively. Cells that belonged to the same cell lineage were marked with the same color (Supplementary Fig. 7a and b), and time-lapse changes of cells in each cell lineage are shown in Supplementary Video 1. The relationships of SNA-I$^+$ cells with their progenitor and progeny cells were determined and analyzed using cell lineage maps (Supplementary Fig. 8). Six of 420 cell lineages (1.4%) contained SNA-I$^+$ progeny (Control-22, 37, 57, 61, 366, and 369 in Supplementary Fig. 8), suggesting that computerized single-cell lineage tracking analysis is capable of characterizing minor cell populations at the single-cell level.

Growth of SNA-I$^+$ cells.

We further investigated the characteristics of SNA-I$^-$ and SNA-I$^+$ cell lineages by creating cell lineage databases for SNA-I$^-$ and SNA-I$^+$ cells, and combining the SNA-I$^+$ cell lineage database with other SNA-I$^+$ cell lineages from other imaging videos (Supplementary Fig. 9). The cell growth curve, normalized to 100 cells at time 1, showed continuous growth of HeLa cells regardless of the confluency of the cell culture (Fig. 4b, orange line), consistent with our previous study (*1*). The growth

of cell lineages containing SNA-I$^+$ cells was plotted after normalization of the numbers of cells at time point 1 to 100 cells (Fig. 4b, blue line). The growth rate of progeny cells in the SNA-I$^+$ cell lineages was significantly higher than that of the HeLa cell population (Fig. 4c, SNA-I$^-$ vs. SNA-I$^+$). We further characterized the cell lineages containing SNA-I$^+$ cells and found that some of these lineages contained both SNA-I$^+$ and SNA-I$^-$ cells (e.g., Supplementary Fig. 7d and Supplementary Fig. 9). The example in Supplementary Fig. 7d suggests that SNA-I$^+$ (red lines) or SNA-I$^-$ (blue lines) cells were produced following cell division (indicated by arrows), raising the possibility that either SNA-I$^-$ cells were produced from SNA-I$^+$ progenitor cells (Supplementary Fig. 7e), or SNA-I$^+$ cells were produced from SNA-I$^-$ progenitor cells (Supplementary Fig. 7f). Because the growth rate of SNA-I$^+$ cells was higher than that of the overall HeLa population (Fig. 4b), the percentage of SNA-I$^+$ cells would increase with every cell passage if SNA-I$^+$ cells were produced from SNA-I$^-$ progenitor cells (Supplementary Fig. 7f). However, the percentage of SNA-I$^+$ cells remained constant at 1%–2% in HeLa cell populations, suggesting that SNA-I$^-$ cells were produced from SNA-I$^+$ cells (Supplementary Fig. 7e).

Cell doubling time of SNA-I$^+$ cells.

The mean cell doubling time was 1532 min (Fig. 4d, Gaussian distribution), with no significant difference between the SNA-I$^-$ and SNA-I$^+$ cell lineages (Fig. 4e), although the growth rate of cells in the SNA-I$^+$ cell lineage was faster than that of the SNA-I$^-$ cell population (Fig. 4b). Notably, the SNA-I$^-$ cell population contained cells with doubling times >2320 min (2.7%, Fig. 4f, data ranges), whereas the SNA-I$^+$ cell populations contained no such cells. It is therefore possible that the SNA-I$^+$ cell population was composed of cells, most of which had similarly fast-growing abilities. To gain further insight into the characteristics of cells in the SNA-I$^+$ cell lineage, we grouped the tracked cells based on the numbers of progeny produced from a progenitor cell at time 693 (i.e. 6930 min).

Maintenance of the cell doubling rate of SNA-I$^+$ cells.

All tracked HeLa cell lineages, including SNA-I$^+$ cell lineages, were grouped according to the numbers of progeny that were present at 6930 min: A; 0–4, B; 5–9, C; 10–14, D; 15–19, E; 20–24; and F ≥ 25 progeny per cell lineage. The cell doubling time for each group of cells is shown in Fig. 5a. In the case of SNA-I$^-$ cell lineages, 41% of the lineages belonged to Groups D–F (Fig. 5a and b), which had significantly shorter mean cell doubling times relative to the other cell groups, i.e., Groups A–E. In the case of lineages containing SNA-I$^+$ cells, 85% of SNA-I$^+$ cells belonged to Groups D–F and none belonged to Groups A or B (Fig. 5a and b), suggesting that SNA-I$^+$ cells were mainly composed of cells with short cell doubling times. Among the SNA-I$^-$ cell population, cell death was observed less frequently in Group C–F cells compared with Group A and B cells, and was similarly less frequent in Groups C–F among SNA-I$^+$ cells (Fig. 5c), indicating that most SNA-I$^+$ cells had similar characteristics to Groups D–F among SNA-I$^-$ cells. However, if SNA-I$^+$ cells represent stem-like cells, they would be expected to have distinct characteristics from SNA-I$^-$ cells. We thus hypothesized that progeny of SNA-I$^+$ cells retained the cell doubling rate of their parental cells relative to SNA-I$^-$ cells, which may lose their reproductive ability over time, as indicated by the fact that most HeLa cells are mortal according to previous single cell lineage tracking analyses (*1*). As shown in Fig. 5d, among SNA-I$^-$ cells, the cell doubling time of daughter cells (DA) was shorter than that of granddaughter cells (GD), while the GD cells showed similar cell doubling time to great-granddaughter cells (GGD). We found a similar tendency in SNA-I$^+$ cells. We therefore tested the above hypothesis by analyzing the time from the bipolar cell division that produced the GD cells until the time of the bipolar cell division that produced the GGD cells (combined cell doubling time of GD+GGD). As illustrated in Fig. 5e, we first determined the shortest cell doubling time of GD+GDD cells (red arrow) among the progeny of a cell lineage, and then calculated the difference between the shortest combined cell doubling time and other combined cell doubling times (Fig. 5e, difference between the blue and green lines). We expected that stem-like cells would show a smaller difference than non-stem cells, if the stem-like cells retained their rate of cell doubling. In support of this hypothesis, the analysis revealed that the difference for SNA-I$^+$

cells was significantly smaller than that for SNA-I^- cells (Fig. 5f). Furthermore, the difference for SNA-I^+ cells in group F was significantly smaller than that for Groups C, D, and E cells (Fig. 5g), although no significant difference was found between SNA-I^- of Group F cells and SNA-I^+ of Group F cells. These results suggest that the rate of cell doubling was retained among progeny of SNA-I^+ cells compared with SNA-I^- of Group C, D, and E cells.

Reproductive ability of SNA-I^+ cells in group F.

To confirm that SNA-I^+ cells retained their reproductive ability relative to SNA-I^- cells, we analyzed the numbers of progeny cells produced from progenitor and GD cells. To perform the analysis, the cell cycle was synchronized *in silico* (*1*), and we determined the numbers of progeny cells produced from a progenitor cell in each cell lineage at 3000 min from the time point when a progenitor cell produced a DA cell, and the numbers of progeny cells produced by each of their GD cells at 3200 min from the time point when the GD cells produced GGD cells (Fig. 6a). For the latter calculation, the period was adjusted to 3200 min by 1.067, considering the difference in growth rates of DA and GD cells (Fig. 5d). If GD cells retained a similar reproductive ability to their progenitor cell, both progenitor and GD cells would be expected to produce the same number of progeny as their progenitor cell. For example, as shown in Fig. 6b, 38.2% of SNA-I^- of Group D progenitor cells (SNA-I^-, D, Progenitor) produced 7–8 progeny. In contrast, only 2.7% of SNA-I^- of Group D DG cells (SNA-I^-, D, GD) produced 7–8 progeny, suggesting that SNA-I^- of Group D GD cells have lost their reproductive ability. We found similar tendencies in SNA-I^- of Groups A–E cells, suggesting that each group of cells tends to lose their reproductive ability during growth. With regard to SNA-I^- of Group F cells, both progenitor and GD cells produced 7–8 progeny more frequently than the other groups, and SNA-I^+ of Group F cells (SNA-I^+, F, Progenitor and GD) showed similar tendencies to Group F cells (SNA-I^-, F, Progenitor and GD), while SNA-I^+ of Group F GD cells retained their reproductive ability (SNA-I^+, F, GD, d (7–8)) better than SNA-I^- of Group F GD cells (SNA-I^-, F, GD, d (7–8)). SNA-I^+ of Group F cells thus showed

similar characteristics to SNA-I$^-$ of Group F cells, except that SNA-I$^+$ expressed α2-6Sia and tended to retain the ability to produce 7–8 progeny.

Simulation of HeLa cell growth.

These results so far suggest that: 1) SNA-I$^-$ cells are produced from SNA-I$^+$ cells (Supplementary Fig. 7e); 2) the progeny of SNA-I$^+$ of Group F cells retained their rate of cell doubling more than other groups of cells, except SNA-I$^-$ of Group F cells (Fig. 5f and g); and 3) SNA-I$^+$ of Group F cells retained their reproductive ability better than SNA-I$^-$ of Groups A–F cells. We therefore assumed that SNA-I$^+$ cells acted as stem-like cells in HeLa cell culture. If the assumption was correct, HeLa cell population homeostasis would be maintained by the self-renewal of SNA-I$^+$ cells and the production of SNA-I$^-$ progeny, which are annihilated by losing their reproductive ability. To test this, we performed a simulation assuming that SNA-I$^+$ cells were HeLa stem-like cells (Table 1) and showed that 1.4% of cell lineages contained SNA-I$^+$ progeny cells (Table 1, Cells (%), SNA-I$^+$ cell lineages, Control-22, 37, 57, 61, 366, and 369 in Supplementary Fig. 8, 6 out of 420 lineages), of which progenitors were expected to be SNA-I$^+$ cells. The total number of progeny of the SNA-I$^+$ cell lineages at the end of imaging was 134 cells (Table 1, number of progeny), while those progeny cells were composed of both SNA-I$^+$ and SNA-I$^-$ cells (e.g. see Supplementary Fig. 7d). In those lineages, 83 cells were SNA-I$^+$ and this corresponded to 1.6% of total progeny cells (Table 1, number of SNA-I$^+$ progeny, 83/5910), which was close to the percentage of the initial SNA-I$^+$ progenitor cells. This suggests that the proportion of SNA-I$^+$ cells within the HeLa cell population is maintained at a similar level during the growth of HeLa cells. Given that four out of six SNA-I$^+$ cell lineages belonged to Group F, we performed calculations assuming that SNA-I$^-$ cells produced from SNA-I$^+$ progenitor cells (Table 1, number of SNA-I$^-$ progeny, 134−83=51 cells) also belonged to SNA-I$^-$ of Group F cells. The number of progeny of SNA-I$^-$ of Group F cells at the final time point (time point 693) was 1504 cells (Table 1, number of progeny). Thus, adding the 51 cells derived from SNA-I$^+$ progenitors gave an adjusted number of

progeny of SNA-I^- of Group F cells of 1555 (Table 1, Adjusted number of progeny, 1504+51=1555). The percentage of SNA-I^- of Group F cells at time 1 was 13% (Table 1, Cells (%)), and assuming that the percentage of this group of cells at the end of imaging (time point 693) was also 13%, the expected number of Group F cells at time 693 was 674 cells (Table 1, Expected number of progeny, 5910×0.13 = 674 cells). The difference between the adjusted number of progeny (1555 cells) and the expected number of progeny (674 cells) (Table 1, Difference, 1555−647=881 cells) was thus the number of cells that no longer belonged to SNA-I^- of Group F, and we assumed that 961 cells, corresponding to 16.9% of the total cell population, became Group E cells. We performed this calculation for Groups E followed by D, C, B and A cells. Because Group A cells terminated cell growth, the "Difference" for Group A cells was expected to be 0. We calculated the "Difference" as -32 cells, which was <1% of the expected number of cells. This simulation thus suggests that the composition of HeLa cells could be accounted for by the growth of 1.4%–1.6% of SNA-I^+ cells and their progeny, which lost their reproductive ability over time. We therefore concluded that SNA-I^+ cells are stem-like cells in HeLa cell cultures.

Discussion

In the present study, we developed a computerized single-cell lineage tracking analysis system and used it to identify SNA-I^+ cells as stem cell-like cells responsible for maintaining the HeLa cell population. DIC-based imaging is a highly versatile imaging method and can be performed without any prior treatment of cells, such as transfection of an expression construct for a fluorescent-tagged protein. However, the development of single-cell tracking analysis systems for DIC-based image is highly challenging (*5*) because of the need to overcome various difficulties, such as creating optimal images for processing, performing accurate image segmentation, and tracking individual cells, irrespective of cell type and confluency and image magnification. For example, cell biology experiments are often performed at confluencies >70%, but currently available mathematical and probability-based approaches may not be optimal for image segmentation and tracking of cells using images of confluent cultured cells. We therefore developed a new approach for segmenting images using SAE and tracking of segmented areas by *Automatic Object Tracking* to track highly confluent cells at the single-cell level by mimicking cell tracking approaches used by the human brain, supported by a series of software programs to coordinately assist live cell video making, database management, and data analysis. Ensuring the stability of the microscope for long-term live cell imaging is also a critical factor for the development of single-cell tracking analysis systems, and we therefore built a microscope suitable for continuous imaging without accidental interruptions. To stabilize the image-acquisition computer (Computer 1), image files created by the computer were transferred to a data-processing computer (Computer 2) to harmonize the performance of the microscope with a series of image processing software programs. In summary, we created a computerized single-cell lineage tracking analysis system able to provide accurate single-cell tracking data for cultured cells. Although we built a custom made-to-order microscope for this study, a suitable microscope can be built using commercially available equipment. This computerized single-cell lineage tracking analysis system can thus be used as a routine cell biological tool to characterize cultured cells at the single-cell level and to open up new avenues for

cell biological studies by revealing spatiotemporal alterations and responses of individual cells to external substances.

In this study, we demonstrated that SNA-I$^+$ cells acted as stem-like cells in HeLa cell cultures. We previously suggested that HeLa cell populations contained well-growing sub-populations, which we referred to as putative cancer stem cells, although the identity of the cells had not then been clarified (*1*). SNA-I$^+$ cells were among the potential cancer stem cells, and the current results confirmed their role as stem cell-like HeLa cells, revealing unique characteristics of the HeLa cell line. Several studies suggested that one cancer stem cell phenotype involves the modification of surface glycans with the α2-6Sia epitope, which is recognized by the plant lectin, SNA-I (*8-11*). This SNA-I epitope is associated with enhanced proliferation activity of pancreatic cancer cells. Other studies have suggested that inducible pluripotent stem cells and adult stem cells, but not somatic or differentiated cells, are also SNA-I$^+$ cells (*11, 20*). It is therefore plausible that the expression of surface glycan modified with the α2-6Sia epitope, recognized by SNA-I, has a critical role in maintaining stemness. Similar to HeLa cells, A549 lung carcinoma cells also contain a well-growing cell population (*23*) and established embryonic fibroblasts were composed of fast- and slow-growing cells (*24*). Furthermore, many established cell lines are composed of individual cells that produce colonies of various sizes, implying that these cells have different reproductive abilities. Thus, although established cultured cell lines have been used for various investigations based on the assumption that they are composed of cells that share similar phenotypic characteristics, the current and earlier observations suggest that many established cell lines may grow in a similar manner to HeLa cells, implying that they may also contain stem-cell like cells. It is therefore necessary to evaluate cell-to-cell heterogeneity empirically in other cell lines.

Finally, our results suggest that HeLa cells may provide an *in vitro* culture model for stem cell growth. Cell population homeostasis *in vivo* is maintained by self-renewal of stem cells and annihilation of differentiated cells. Although cancer cells disturb this homeostasis, cancer cell lines

may become established because an isolated clone happened to be a stem-like cell, and growth of its progeny was balanced between self-renewal and annihilation of some of the progeny, as simulated in this study. Established cancer cell lines may thus be used as new research materials to investigate the maintenance of cell population homeostasis by stem-like cells and its potential disruption by anti-cancer therapies.

Materials and Methods

Cell culture and cell plating.

HeLa and C2C12 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS) in a humidified 5% CO_2 atmosphere. C2C12 cells were differentiated by replacing the medium with DMEM containing 1% horse serum and ITS liquid media supplement (Sigma) as previously described (*25*). Cells were plated onto a cover glass Lab-TekII 8-well chamber by placing 50 μl of cell suspension (approximately 3500 cells) in the center of each well and allowing the cells to attach to the cover glass surface, followed by the addition of 0.75 ml of culture medium to each well. The Lab-TekII 8-well chamber was placed on a microscope stage 24 h after plating.

Microscope.

We built a customized microscope (Quorum Technologies, Ontario, Canada) using an Olympus IX81 microscope frame. We used NIR-DIC imaging to minimize phototoxity and because NIR-DIC is less affected by light distortion caused by plastic partitions and shape changes in the surface of the medium. We combined the NIR-DIC imaging with fluorescent imaging (live and fixed sample) by creating two light paths in the microscope (Supplementary Fig. 1a). For NIR-DIC imaging, NIR light was created by light-emitting diodes (LEDs), passed through a polarizer, Nomarski prism, and condenser, and then used to illuminate the cells. The returned light then passed through an objective lens and Nomarski prism and was brought to the first charge-coupled device (CCD) camera (Camera 1, Hamamatsu photonics, Image EM, 512×512 pixels). For fluorescent imaging, laser light reached the cells that were stained with a fluorescence-conjugated antibody or protein by passing through a Nipkow disk (Yokokawa-Quorum Technologies) and an objective lens to excite a fluorophore. The emitted light was then brought to the Nipkow disk and captured using the second CCD camera (Supplementary Fig. 1, Camera 2, Hamamatsu, Image EM, 512×512 pixels).

The computerized single-cell lineage tracking analysis system allows the use of ×10–×40 objectives. In this study, we used an Olympus ×10 dry objective (UPlanSApo, 10×/0.40 NA, α/0.17/FN26.5) or ×20 dry objective (UPlanSApo, 20×/0.75 NA, α/0.17/FN2G.5) and images were magnified using a ×1.5 coupler (Quorum Technologies) inserted into the light path to the CCD cameras, to create images equivalent to ×15 and ×30 magnifications, respectively. Piezo stages with relative measurement often cause image shifts during 1–3 weeks of imaging, and we therefore used a precision piezo XY stage with absolute measurement. An environmental chamber (Live Cell Instrument, Korea) was set onto the piezo stage, and the microscope and piezo XY stage was controlled using Metamorph (Quorum Technologies), installed on a Windows computer (Fig. 1, Computer 1). Images were acquired using Metamorph's multi-dimensional acquisition mode (MDA) with 34 ms exposure time for NIR-DIC imaging, and the exposure time for fluorescence imaging was adjusted based on the fluorescence intensity. Typically, 20–80 z-planes were acquired every 1 μm for NIR-DIC imaging and the acquired z-plane NIR-DIC images captured by Camera 1 were saved as 512×512 pixel multi-layer TIFF files. For fluorescence imaging, z-plane NIR-DIC and fluorescence images (up to three channels per image acquisition) created by CCD cameras 1 and 2, respectively, were merged using a macro program in Metamorph, which created 512×1024 pixel multi-layer TIFF files.

Computer 1 (Windows computer) was linked via an Ethernet cable to a Macintosh computer (Computer 2), which created live cell videos (Fig. 1). We previously used a commercially available microscope and performed live cell imaging on the same Windows computer used to control the microscope and stage and save cell images (*1*); however, this resulted in frequent interruptions to the imaging and processing due to crashes of either the Windows computer or the Volocity software used to control the operation of the microscope throughout the 1–3 weeks of continuous imaging. In the current study, we therefore designed the microscope to minimize the risk of termination due to crashes. The performance of the Windows computer was affected by the accumulation of a large number of files

on its hard drive, and defragmentation could not be performed during image acquisition. We therefore stabilized the Windows computer by immediately transferring the image files to Computer 2 (Macintosh) using *File Transfer* and deleting the image files from Computer 1. We also noted previously that electrical noise created by tungsten or halogen light sources could affect the behavior of the software controlling the microscope, and we therefore replaced the light sources with LEDs. Finally, the microscope and all other related apparatus and computers were attached to uninterruptible power supplies to stabilize the electrical current.

Culture conditions on the microscope stage.

Cells were cultured in an environmental chamber at 37°C and 65% relative humidity. The concentration of CO_2 was maintained at 7.5% because lower CO_2 concentrations increased the pH of the medium in the cover glass Lab-TekII 8-well chamber. An NIR-DIC optimal glass lid (Live Cell Instruments) was placed on the chamber to minimize evaporation of the medium. The typical medium evaporation rate was 10 μl/24 h, and we therefore added 700–800 μl medium per well, which allowed the cells to be cultured for at least 1 week without changing the medium. We avoided exposing the cells to the air during changes of medium, because exposure to the air for even a few seconds could cause detectable cytotoxicity. We monitored the pH of the medium closely during long-term live cell imaging by including phenol red in the medium, which did not interfere with NIR-DIC imaging. Finally, it should be noted that medium containing insulin may be cytotoxic if exposed to white light, and we therefore confirmed that NIR-DIC did not induce insulin-related cytotoxicity.

Developed software.

Image files created by Metamorph were processed using a series of custom software (C/C++/Objective-C for MacOS): *Image Processing Controller*, *File Transfer*, *Map*, *Name Assignment*, *Focus Image Selection*, *Contrast Set*, *Data Backup*, *Image Viewer*, *Outline Drawing* (for image

segmentation), *Object Tracking Controller*, *Automatic Object Tracking*, *Data Analysis*, and *File Converter* (Supplementary Table 1). *Image Processing Controller* linked to all the other programs to control image processing, but each software program was also designed to operate independently to reduce the risk of failure of the entire system due to a crash of one software program. *Image Processing Controller* monitored the activities of the other software programs and re-launched a software if it crashed. These software programs were designed to perform long-term live cell imaging, image data segmentation, and object tracking in real time. Although Computer 2 (Fig. 1) was nominally a single computer (Fig. 1), currently available Macintosh computers did not have sufficient capacity to achieve all the processing in real time, and *Outline Drawing*, *Object Tracking Controller*, *Auto Object Tracking*, and *Data Analysis* were therefore installed on a third (Macintosh) computer, linked to the second computer via an Ethernet cable.

Setting FOVs and adjustment of focus.

Cells cultured under different conditions were monitored simultaneously using an 8-well chambered slide, with 2D image acquisition arrays created in each well to cover the area of interest with multiple FOVs (Supplementary Fig. 1b) (*1*). Setting the FOVs for the 2D image acquisition arrays and focus adjustment were performed using *Map* in conjunction with Metamorph. Because Metamorph software does not have an option to display the objective position on the monitor, *Map* was designed to coordinate and share the information on the xyz positions between Metamorph in Computer 1 and our software in Computer 2. To initialize this coordination, the objective was moved to the left corner of Well 2 and the position was registered in *Map*, and the outline of the cover glass Lab-TekII 8-well chamber was then drawn on the screen (Supplementary Fig. 2a). The objective was moved to an FOV of interest in a well, and the surrounding images (e.g. 5×5 FOV dimension) were acquired using Metamorph. The xyz position for each FOV was saved on Computer 1, and *Map* on Computer 2 then used these data to display the xy positions as light gray boxes (Supplementary Fig. 2a). The 2D image

acquisition arrays of interest were then set by selecting a corner position of an array from one of the FOVs, and displayed as red boxes (Supplementary Fig. 2a). We then determined the optimal objective z position for each FOV, and created an MDA file containing the information on the xy positions and the optimal z positions. The file was sent to Computer 1 and uploaded to Metamorph, which then performed MDA image acquisition to create images for the corresponding FOVs.

Irregular z shifts occasionally occurred during long-term live cell imaging. To continuously obtain images with optimal NIR-DIC focus, we captured 20–80 z-planes at optimal z positions at the starting point. The images were transferred to Computer 2 as multi-layer TIFF files and were read by *Map*, and images of each plane were displayed to find the optical focal plane. The objective z position was adjusted such that the focus position was in the middle of the z-planes, e.g., if 30 z-planes were taken, then the objective z position was adjusted so that the 15th z-plane was the focal plane. After determining the z position for all wells, xyz positions for all FOVs were written to an MDA file by *Map* and sent to Computer 1. The file was uploaded to Metamorph to perform MDA image acquisition. Typically, a total of 300 FOVs could be obtained by taking images using 34 msec exposure and 30 z-plane during a single image acquisition cycle within 10 min.

Starting image acquisition.

Image quality varied between acquisitions, and changed over time due to light distortion caused by changes in the shape of the medium surface and cell density, and the emergence of cell debris. Mechanical factors, e.g., shadows created by dust, also affected image quality. To automatically generate optimal images for segmentation and object tracking, we set the image acquisition conditions before commencing fully automated image acquisition and live cell video creation. The first round of image acquisition using Metamorph commenced after creating the MDA and uploading it to Computer 1. *File Transfer* (Computer 2) was launched during image acquisition to monitor the acquisition progress, and the multi-layer TIFF files (512×512 pixels) created by Metamorph were transferred to

Computer 2 when all FOV acquisition was completed, and the files were then deleted from Computer 1. For 512×1024-pixel image files, which contained both NIR-DIC and fluorescence images, the images were split into individual NIR-DIC and fluorescence image files (512×512 pixels). The each multi-layer TIFF file was assigned a new file name by *Name Assignment* to allow for file archiving. The image files were processed using *Focus Image Selection* (Supplementary Fig. 2b) as follows. The each multi-layer TIFF file contained unfocused and focused z-plane optical slices (Supplementary Fig. 2c). *Focal Image Selection* created a single-layer TIFF file by either automatically selecting a particular focal plane from the multi-layer TIFF or creating an all-in-focus image, which was reconstructed by selecting the optimal grayscale value (0–255) from the z-axis data for each pixel (Supplementary Fig. 2b). In the case that a particular focal plane was selected, the Sobel edge extraction method and the standard deviation of the pixel values was used. In the case that an all-in-focus image was created, optimal grayscale value was determined by analyzing values of neighboring pixels of each pixel from each z-plane. For fluorescence images, an extended-focus image (stacked image) was created by selecting the highest value pixel among the z stack images, and/or by calculating the average pixel values across all the z-planes. This software assigned the processing type for each FOV and recorded the actual conversion options used for each FOV for each time point (Supplementary Fig. 2d). Resulting images were 8-bit grayscale single-layer TIFF files. Contrast adjustment and background correction were then performed using *Contrast Set* (Supplementary Fig. 3a). Because the background pattern varied for each FOV, we created customized background patterns using data from single-layer TIFF files. We attempted to use images obtained with an empty chamber as the background pattern, but this was insufficient to optimize image quality for analysis due to subtle but critical optical distortion differences caused by the medium in the chamber well or the plastic structures of the well. We therefore used images from the initial time point to create customized background patterns (Supplementary Fig. 3b). The original image (512×512 pixels) was converted into 64×64 blocks (8×8 pixels each), and then an object (e.g. Supplementary Fig. 3b) was filled by pixel values right and left,

or up and down from the object. All objects in the image were filled following the same process to create the customized backgrounds (Supplementary Fig. 3b). Each background pattern was then applied to corresponding FOVs using *Contrast Set* (Supplementary Fig. 3c). Because even a small CCD camera tilt created a significant FOV position shift, *Contrast Set* also adjusted the position of the FOV as required. Position shifts between NIR-DIC and fluorescence images were also adjusted. Finally, stitched images were created using *Contrast Set* after contrast adjustment, background correction, and fine adjustment of position for each FOV (Supplementary Fig. 3a). Supplementary Fig. 3d shows a sample stitched image before background correction. Black edges and unevenness were normalized after correction (Supplementary Fig. 3e). During the process of background correction, blank parts without cells were filled by setting each pixel value to 100 on a grayscale of 0–255, and values of 95–99 and 101–105 converted to 94 and 106, respectively.

Fully automated long-term live cell imaging video creation.

In the case of HeLa cells, 5×5 image acquisition arrays (25 FOVs) were created in each well and images were acquired every 10 min. Automated live cell video creation commenced by launching *File Transfer*, *Name Assignment*, *Focus Image Selection*, *Contrast Set*, and *Data Backup,* under the control of *Image Processing Controller*. The progression of image acquisition and transfer of the resulting multi-layer TIFF files from Computer 1 to Computer 2 were monitored by *File Transfer*. Appropriate file names were assigned by *Name Assignment*, and the files were processed by *Focus Image Selection* and *Contrast Set*, saved into a temporary folder until stitched images were created, and the relevant files were then moved to the folder holding the created images. The multi-layer TIFF files were copied to external hard drives and deleted from Computer 2 using *Data Backup*. We confirmed that long-term live cell imaging could continue for at least 4 weeks for 5×5 arrays created in eight wells with images acquired every 10 min. *File Transfer*, *Name Assignment*, *Focus Image Selection*, *Contrast Set*, and *Data Backup* were launched individually to start automated live cell video creation and the live cell

video archive was created automatically. When a focal position shift occurred, the z position was readjusted using *Map*, creating a new MDA file that was sent to Computer 1. MDA image acquisition was temporarily placed on hold, the new MDA file was uploaded, and MDA image acquisition restarted.

Image segmentation.

Image segmentation was performed with *Outline Drawing* using SAE (see Results, "Segmentation of NIR-DIC image").

Cell lineage database.

Object tracking was performed using *Object Tracking Controller* (see Results; "Controlling object tracking", and "Automated object tracking"). Cell tracking results were entered into the cell lineage database, which contained data regarding time points, cell lineage numbers, cell numbers, x and y positions of the cells, types of events that occurred in the cells, sizes, average brightness, and gravity centers of the areas, information on the edge circles, fluorescence intensities of the cells, types of fluorophore, relationships of parent cells and their progeny, and segmented area numbers. This database comprised six data arrays, and each data point was stored in at least two arrays, such that if a data array became corrupted, it could be recovered using data stored in the other arrays. *Object Tracking Controller* verified and repaired the cell lineage database structure as required.

Automatic and interactive setting of cellular events.

Object Tracking Controller and *Automatic Object Tracking* allowed the interactive and automatic detection of cell division (bipolar, tripolar, and tetrapolar), incomplete cell division, mitosis, cell fusion, cell death, and loss of cells (Supplementary Fig. 5a). The interactive tracking was performed by displaying live cell imaging videos using *Object Tracking Controller* with tools that could change the

shape of edge circles by cutting or merging circles, and remove, add, or insert events. The automatic detection of cell divisions was carried out by *Automatic Object Tracking*, which first detected mitosis using data for various mitotic shapes recorded in the Mitosis Pattern Library (Supplementary Fig. 5a). This library included the shapes of mitotic cells observed during object tracking. *Automatic Object Tracking* defined bipolar, tripolar, or tetrapolar cell divisions when two, three, or four bright areas were identified, respectively, as possible tracking targets after detection of mitosis. Cell death was identified when a bright area, which constituted debris from cell death, was detected for >50 time points (500 min). When one area corresponding to a cell was attached to an area corresponding to another cell for 50 time points, it was temporally tagged as a possible cell fusion event. If cellular events other than bipolar cell division were identified, information on the cell was entered into the *Check* list for data verification. The accuracy of automatic detection of those events varied depended on cell density and type. For example, *Automatic Object Tracking* detected mitosis automatically (Supplementary Video 3, video Frame 1 vs. Frame 2) in cells that formed a typical round shape during mitosis, whereas mitosis was difficult to detect in cells, which do not form a typical round shape during mitosis, requiring marking for mitosis and cell division using interactive tools. It is therefore essential that cell tracking software includes optimal interactive functions to adjust the results of automatic object tracking to ensure efficient and accurate single-cell tracking.

Data analysis software.

Data Analysis software can hold multiple databases of cell lineages simultaneously to allow various analyses, including cell growth, cell division time, cellular event evaluation and comparison, and cell lineage mapping (Supplementary Fig. 5b). Other analyses, e.g. motility, and matching cell lineage information with end-point staining (fluorescence images) can also be added. *Data Analysis* was designed as a platform to allow unlimited types of analyses depending on the particular requirements.

Staining of cells with SNA-I-TRITC and position adjustment of fluorescence images.

We performed fluorescent imaging at the end of live cell imaging. Before removing the Lab-TekII 8-well chamber from the microscope stage, a snapshot (Snapshot 1) was taken using *Map* to record a unique object, e.g. a unique cell shape (Supplementary Fig. 6a and b, Snapshot 1). The chamber was then removed from the stage, and the cells were washed three times with phosphate-buffered saline (PBS), fixed with 3.7% paraformaldehyde for 15 min at room temperature, and washed three further times with PBS. Cells were then treated with Carbo-Free Blocking Solution (×1) (Vector) for 1 h at room temperature and incubated with TRITC-labelled SNA-I (E-Y Laboratories) diluted with Carbo-Free Blocking Solution (×1) to 50 μg/ml for 1h at 4°C. Subsequently, cells were washed with PBS three times and treated with DAPI by diluting two drops of NucBlue Fixed Cell ReadyProbes Reagent with 1 ml Carbo-Free Blocking Solution (×1) for 15 min at room temperature, followed by three further washes with PBS. The chamber was put back on the microscope stage and the position of the chamber was adjusted to include the unique object in the FOV. A second snapshot (Snapshot 2) was then taken (Supplementary Fig. 6b, Snapshot 2) and the identical object in Snapshots 1 and 2 was marked using *Map* (Supplementary Fig. 6b). *Map* then calculated the difference between the x and y positions in the two snapshots. Position shifts due to this adjustment were generally within two pixels (Supplementary Fig. 6c). A new MDA file was created taking the position shift into account and uploaded to Metamorph. Image acquisition commenced after position and focus verification for each FOV.

Post-acquisition processing of images from backup files.

If the image quality of a live cell video was not optimal, we performed post-acquisition image processing. If the image focus was sub-optimal, we processed back-up multi-layer TIFF files created by Metamorph, and if the contrast and/or background correction was sub-optimal, post-acquisition image processing of the backup single-layer TIFF files was performed to reset the brightness and contrast, and custom backgrounds were reassigned to corresponding FOVs using *Contrast Set*.

Fine-tuning of live cell videos.

Live cell videos were reviewed using *Image Viewer*. Image shifts occasionally occurred during long-term live cell imaging, and image positions were adjusted using *Image Viewer*.

Data analysis.

Cell growth curves, cell doubling times of individual cells, grouping of cells based on the numbers of progeny produced from a progenitor cell, and the duration of combined cell doubling times of GD and GGD were calculated using *Data Analysis*. Data were exported and analyzed using Excel and/or Prism 7 for statistical analyses and graphical visualizations.

Staining of cells with antibody against a stem cell marker.

HeLa cells were fixed, washed with PBS, and stained with SNA-I-TRITC, as described above ("Staining of cells with SNA-I-TRITC and position adjustment of fluorescence images"). Cells were stained with CD133 (Developmental Studies Hybridoma Bank) diluted with Carbo-Free Blocking Solution (×1) to 5 μg/ml for 30 min at room temperature or TAG-72 (B72.3) at 1:50 dilution with Carbo-Free Blocking Solution (×1) for 1 h at room temperature. The cells were then washed three times with PBS and incubated with secondary antibody (Invitrogen, goat anti-mouse IgG Alexa Fluor 488 diluted 1:1000) for 1 h at room temperature. Cells stained with FITC-rBC2LCN (Wako) were incubated with FITC-labelled rBC2LCN at 1:100 dilution with Carbo-Free Blocking Solution (×1) for 30 min at room temperature. Cells were stained with DAPI and washed three times with PBS, as described above ("Staining of cells with SNA-I-TRITC and position adjustment of fluorescence images"). Fluorescence imaging was performed to evaluate expression of CD133, TAG-72 (B72.3), α1-2Fucose modification (rBC2LCN), and α2-6Sia (SNA-I).

Statistical analyses.

Data were analyzed using Welch's *t*-test (unpaired and two-tailed) or one-way ANOVA (Tukey's multiple comparison test) using Prism 7.

Acknowledgements

The authors acknowledge the Bioimaging Platform at the Centre de Recherche Centre Hospitalier Universitaire (CHU) de Quebec. The mAb against CD133 was obtained from the Developmental Studies Hybridoma Bank, created by the U.S. National Institutes of Health, Institute of Child Health and Human Development, and maintained at the Department of Biology, University of Iowa (Iowa City, IA, USA).

References

1. S. Sato, A. Rancourt, Y. Sato, M. S. Satoh, Single-cell lineage tracking analysis reveals that an established cell line comprises putative cancer stem cells and their heterogeneous progeny. *Sci Rep* **6**, 23328 (2016).
2. D. C. Roberts, G. E. Cole, The viability and capacity for further division, following tripolar mitosis, of cells of a murine ascites carcinoma in vitro. *Int J Cancer* **5**, 238-243 (1970).
3. O. Al-Kofahi *et al*., Automated cell lineage construction: a rapid method to analyze clonal development established with murine neural progenitor cells. *Cell Cycle* **5**, 327-335 (2006).
4. R. Bise, T. Kanade, Z. Yin, S. I. Huh, Automatic cell tracking applied to analysis of cell migration in wound healing assay. *Conf Proc IEEE Eng Med Biol Soc* **2011**, 6174-6179 (2011).
5. O. Hilsenbeck *et al*., Software tools for single-cell tracking and quantification of cellular and molecular properties. *Nat Biotechnol* **34**, 703-706 (2016).
6. M. Maska *et al*., A benchmark for comparison of cell tracking algorithms. *Bioinformatics* **30**, 1609-1617 (2014).
7. S. Skylaki, O. Hilsenbeck, T. Schroeder, Challenges in long-term imaging and quantification of single-cell dynamics. *Nat Biotechnol* **34**, 1137-1144 (2016).
8. C. M. Britain, K. A. Dorsett, S. L. Bellis, The Glycosyltransferase ST6Gal-I Protects Tumor Cells against Serum Growth Factor Withdrawal by Enhancing Survival Signaling and Proliferative Potential. *J Biol Chem* **292**, 4663-4673 (2017).
9. H. Cui *et al*., The glycosyltransferase ST6Gal-I is enriched in cancer stem-like cells in colorectal carcinoma and contributes to their chemo-resistance. *Clin Transl Oncol*, (2018).
10. M. J. Schultz *et al*., The Tumor-Associated Glycosyltransferase ST6Gal-I Regulates Stem Cell Transcription Factors and Confers a Cancer Stem Cell Phenotype. *Cancer Res* **76**, 3978-3988 (2016).

11. A. F. Swindall *et al*., ST6Gal-I protein expression is upregulated in human epithelial tumors and correlates with stem cell markers in normal tissues and colon cancer cell lines. *Cancer Res* **73**, 2368-2378 (2013).

12. S. P. Povoski, I. S. Hatzaras, C. M. Mojzisik, E. W. Martin, Jr., Oncologic theranostics: recognition of this concept in antigen-directed cancer therapy for colorectal cancer with anti-TAG-72 monoclonal antibodies. *Expert Rev Mol Diagn* **11**, 667-670 (2011).

13. B. E. Rogers *et al*., Intraperitoneal radioimmunotherapy with a humanized anti-TAG-72 (CC49) antibody with a deleted CH2 region. *Cancer Biother Radiopharm* **20**, 502-513 (2005).

14. J. Schmolling, J. Reinsberg, U. Wagner, D. Krebs, Anti-TAG-72 antibody B72.3--immunological and clinical effects in ovarian carcinoma. *Hybridoma* **16**, 53-58 (1997).

15. S. Miraglia *et al*., A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. *Blood* **90**, 5013-5021 (1997).

16. A. Weigmann, D. Corbeil, A. Hellwig, W. B. Huttner, Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc Natl Acad Sci U S A* **94**, 12425-12430 (1997).

17. A. H. Yin *et al*., AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* **90**, 5002-5012 (1997).

18. H. Tateno *et al*., Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J Biol Chem* **286**, 20345-20353 (2011).

19. D. R. Christie, F. M. Shaikh, J. A. t. Lucas, J. A. Lucas, 3rd, S. L. Bellis, ST6Gal-I expression in ovarian cancer cells promotes an invasive phenotype by altering integrin glycosylation and function. *J Ovarian Res* **1**, 3 (2008).

20. K. Hasehira *et al*., Structural and quantitative evidence for dynamic glycome shift on production of induced pluripotent stem cells. *Mol Cell Proteomics* **11**, 1913-1923 (2012).

21. K. Moriwaki *et al*., Combination use of anti-CD133 antibody and SSA lectin can effectively enrich cells with high tumorigenicity. *Cancer Sci* **102**, 1164-1170 (2011).

22. A. F. Swindall, S. L. Bellis, Sialylation of the Fas death receptor by ST6Gal-I provides protection against Fas-mediated apoptosis in colon carcinoma cells. *J Biol Chem* **286**, 22982-22990 (2011).

23. A. Rancourt, S. Sato, M. S. Satoh, Basal level of p53 regulates cell population homeostasis. *BioRxiv* **doi:** https://doi.org/10.1101/319525.

24. S. Guo *et al*., Nonstochastic reprogramming from a privileged somatic cell state. *Cell* **156**, 649-662 (2014).

25. A. Rancourt *et al*., Galectin-3 and N-acetylglucosamine promote myogenesis and improve skeletal muscle function in the mdx model of Duchenne muscular dystrophy. *FASEB J*, fj201701151RRR (2018).

Figure legends

Figure 1. Computerized single-cell lineage tracking analysis system.

The computerized single-cell lineage tracking analysis system included Computer 1 to control the microscope and Computer 2 to perform image processing, single-cell tracking, and data analysis. Commercially available image acquisition software (Metamorph) was installed on Computer 1 to control the microscope and create image files. Computer 2 employed various custom software programs as listed in Supplementary Table 1. *Image Processing Controller* controlled other software programs; *File Transfer* and *Map* communicated between Computers 1 and 2; *File Converter* imported images created by other microscopes; *Name Assignment*, *Focal Image Selection*, and *Contrast Set* produced live cell videos; *Data Backup* controlled file archiving and backup; *Outline Drawing* performed image segmentation and created the object segmentation library; *Movie Viewer* played movies and fine-tuned image quality; *Object Tracking Controller* created the cell lineage database by controlling *Automatic Object Tracking*, and facilitated data verification; and *Data Analysis* provided various options for data analysis. DIC/FL; differential interference contrast/fluorescence.

Figure 2. Image segmentation.

a. Image segmentation was carried out using SAE. Pixels with values above the respective thresholds for TH1–TH4 extracted from the NIR-DIC image and assigned a value of 150 (0–255 grayscale). Areas surrounded by an edge circle (blue line) were segmented. **b.** Connectivity analysis was performed using TH1 to identify connective pixel edges outlined by edge circles (**A**: Edge circles, pink lines). The circles were overlaid onto the NIR-DIC image and lines extended from the pixels within the regions towards the 12, 3, 6, and 9 o'clock directions (**B**: Line extension, red, yellow, light blue and blue regions, respectively, and Line extension (magnified)) to possible cell borders. The ends of each line were linked (**C**: Edge linking, and Edge linking (magnified)) and edge circles were finally created by linking the regions (**D**: Creating new areas: green lines). These edge circles were then overlaid onto the

TH2 image (**E**: green circles) and connected pixel edges that did not overlap with green circles were determined (**E**: pink circles). Pink circles were processed as described for TH1 (**F-H**). Green circles were expanded for 1–4 pixels outside the edge circles (**K** and **L**, and Edge expansion (magnified)). Finally, green and red edge circles were overlaid to detect overlaps (**I**: Overlay), which were then removed (**J**: Removal of overlaps). These processes were repeated for TH3 and TH4 images. Image segmentation for (**c**) low-density HeLa cells, (**d**) high-density HeLa cells, (**e**) C2C12 mouse myoblasts, and (**f**) differentiated C2C12 myotubes.

Figure 3. Object tracking.

a. *Object Tracking Controller* was used to assign cell lineage numbers using the live cell video archive and object segmentation library. Automatic object tracking was performed for each cell in the *Queue* list for 10–20 time points. This step was repeated until time point 100–200. If *Automatic Object Tracking* failed to track a segmented area, the information for that area was entered into the *Check* list. Tracking data were evaluated to determine whether tracked segments represented individual cells using the interactive functions after completing 100–200 time points. Resulting tracking data were stored in the cell lineage database. **b.** *Object Tracking Controller* was used to identify possible target and non-target areas (Tracking target and non-target) that represented a cell (green arrow) and a non-target (yellow arrow). Areas could also be defined as cells using the interactive functions (Tracking target (cell), red arrow). Finally, cell lineage numbers (light orange numbers) were assigned. **c.** Automatic object tracking was performed by dividing or merging segmented areas that surrounded the target segment. Dividing surrounding segments: **A**. Orange characters and numbers indicate segments that were assigned to cells in the time A image. **B.** Cell is marked with light orange circle. The cell indicated by the orange arrow was a tracking target of a cell. **C.** Segments around the cell were circled with red or pink circles, and the segment that was circled with red (indicated by red arrow) represents the segment being tracked. **D.** Black numbers indicate non-assigned segments in the time A+1 image.

Segment patterns at time A+1 differed from time A**. E**. Light orange circle in time A was overlaid to time A+1 image. Segments around the cell are circled with yellow circles. The segments corresponding to the red circle disappeared. One of the yellow segments (yellow circle 2) partly overlapped the target cell (orange circle). **F.** The time A red segment, which represented a target cell (light orange circle), was overlaid on the time A+1 image. The red segment overlapped one of the yellow segments (yellow circle 2). **G.** Surrounding information from the target cell, such as neighboring segments (pink circles 1 and 2) at time A, were overlaid onto time A+1 image. **H.** Because yellow circle 1 was not attached to the red circle and only overlapped pink circle 1, yellow circle 1 was not considered to be a part of the target cell. In contrast, because yellow circle 2 overlapped both pink 2 and the red segment, the segment was divided into two segments (white circles 1 and 2). **I.** The redefined segmentation for time A+1 was drawn by combining the red segment from time A with white segment 2. Merging surrounding segments: **J.** The cell marked with an ivory circle and indicated by the ivory arrow was a tracking target of a cell. The segment circled with light blue represents the segment being tracked. **K.** Time A light blue segment was overlaid on the time A+1 image. The light blue segment contained multiple segments (green circles). **L.** Redefined segmentation for time A+1 was drawn by merging the multiple segments inside the light blue area (blue circle).

Figure 4. Identification, growth, and cell doubling time of SNA-I$^+$ cells in the HeLa cell population.

a. HeLa cells were stained using fluorescence-tagged rBC2LCN, and SNA-I, or anti-CD133 and anti-TAG-72 followed by the Alexa 488-labelled secondary antibody, and visualized by fluorescent imaging. Cells were counterstained using DAPI (blue regions). DIC images are also shown. **b.** The numbers of SNA-I$^-$ and SNA-I$^+$ cells at each time point were determined using the cell lineage database. The initial numbers of cells were normalized to 100. Curves were fitted (R value) using non-linear regression (Prism 7). **c.** The numbers of cells produced from each progenitor cell at the 2000, 4000, and

6930 min time points were calculated and Welch's *t*-tests were performed relative to SNA-I$^-$ cells; $^*p < 0.05$, and $^{****}p < 0.0001$. Means ± standard errors (SEM) are shown. Cell doubling times for SNA-I$^-$ **(d)** and SNA-I$^+$ (**e**) cells are shown. Means and standard deviations were calculated using a Gaussian distribution (Prism 7). **f.** Cell doubling times for SNA-I$^-$ and SNA-I$^+$ cells were analyzed by Welch's *t*-tests; ns, not significant. Means are shown with data ranges.

Figure 5. Analysis of cell doubling time after grouping of cells, cell death, and combined cell division timing differences.

a. SNA-I$^-$ and SNA-I$^+$ progenitor cells and their cell lineages were grouped according to the numbers of progeny cells at 6930 min: A, 0–4; B, 5–9; C, 10–14; D, 15–19; E, 20–24; and F ≥ 25 cells. Welch's *t*-tests were performed relative to Group F for SNA-I$^-$ cells and SNA-I$^+$ cells; ns, not significant, and $^{****}p < 0.0001$. **b.** Composition of each group of cells (%) is shown. **c.** The numbers of cell death events were determined using the cell lineage database. One-way ANOVA (Tukey's multiple comparison test) was performed; ns, not significant, and $^{****}p < 0.0001$. **d.** Cell doubling times for DA, GD, and GGD cells are shown. Welch's *t*-tests were performed relative to DA; $^*p < 0.05$, $^{**}p < 0.01$, and $^{****}p < 0.0001$. **e.** Combined cell doubling times for GD and GGD cells are shown. The shortest combined cell doubling time (GD+GGD) was determined (red arrow) between the bipolar cell division that produced GD cells (pink lines) and bipolar division of a GGD (blue line), and the combined cell doubling times for other cells (black arrows) were then determined (pink to green lines), and the time difference between the green and blue lines was calculated. **f.** The time differences between SNA-I$^-$ and SNA-I$^+$ cells are shown. Welch's *t*-tests were performed; $^{**}p < 0.01$. **g.** The time differences of Group C–F of SNA-I$^-$ and Group F of SNA-I$^+$ cells were analyzed. One-way ANOVA (Tukey's multiple comparison test) was performed. Statistical significance relative to Group F of SNA-I$^-$ or Group F of SNA-I$^+$ cells are shown; ns, not significant, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ and $^{****}p < 0.0001$. **a, c, d, f, and g.** Means ± SEMs are shown

Fig. 6. Analysis of progeny produced from progenitor and GD cells.

a. The cell cycle was synchronized *in silico*. The time point when the progenitor cells divided was normalized as time 1 (blue arrowhead). Blue line indicates 3000 min after the first division, green arrowheads indicate GD cells, red arrowheads indicate the cell division of GD cells, and red line indicates 3200 min after GD cell division. **b.** The numbers of progeny (a, 0–2; b, 3–4; c, 5–6; and d, 7–8 cells) produced from a progenitor cell at 3000 mins and from GD cells at 3200 min from a division were determined. Data are given as a percentage of progeny.

Table 1. HeLa cell growth simulation.

SNA-I$^+$	Cells[2)] (%)	Number of progeny[3)]	Number of SNA-I$^+$ progeny[4)]		Number of SNA-I$^-$ progeny[6)]	Number of SNA-I$^-$ progeny[9)] (% of total cells)
SNA-I$^+$	1.4	134	83 (1.6%)[11)]		51	0.9
SNA-I$^-$[1)]	**Cells[2)] (%)**	**Number of progeny[3)]**	**Expected number of progeny[5)]**	**Adjusted number of progeny[7)]**	**Difference[8)]**	**Difference[10)] (% of total cells)**
F: ≥ 25	13.0	1504	674	51+1504 = 1555	1555-674 = 881	16.9
E: 20–24	10.2	923	519	881+923 = 1804	1804-519 = 1285	24.7
D: 15–19	18.1	1217	943	1285+121 = 2502	2502-943 = 1559	30.0
C: 10–14	16.2	816	840	1559+816 = 2375	2375-840 = 1535	29.4
B: 5–9	18.0	542	934	1535+542 = 2077	2077-934 = 1134	21.8
A: 0–4	24.0	74	1240	1134+74 = 1208	1208-1240 = -32	-0.6

[1)] Progenitor cells were grouped as A–F on the basis of the number of progeny produced at 6930 min.

[2)] The percentage of progenitor cells. We assumed the proportion percentage was maintained during growth.

[3)] The number of progeny at 6930 min.

[4)] The expected number of SNA-I$^+$ progeny at 6930 min.

[5)] The expected number of progeny at 6930 min was e.g. 5210×0.13 = 674 cells, assuming that the composition of each group of cells at 6930 was the same as at the start of imaging.

[6)] The number of SNA-I$^-$ progeny produced from SNA-I$^+$ progenitor cells at 6930 min.

[7)] The adjusted number of progeny for each group was calculated e.g. Adjusted number of progeny (E) = Number of progeny (E) + Difference (F).

[8)] Difference represents the difference between “Adjusted number of progeny” and “Expected number of progeny”.

[9)] The percentage of “Number of SNA-I^- progeny” within the total number of cells at 6930 min.

[10)] The percentage of “Difference” within the total number of cells at 6930 min.

[11)] The percentage of “Number of SNA-I^+ progeny” within the total number of cells at 6930 min.

Supplementary Movie 1. Single HeLa cell tracking.

A typical video for single-cell tracking is shown (time point 1–693). A unique color was assigned to each cell lineage. The total number of tracked cells was 11,640.

Supplementary Movie 2. Single HeLa cell tracking (magnified).

Part of the video from Supplementary video 1 was magnified.

Supplementary Movie 3. Automatic cell tracking video.

Typical automatic cell tracking video is shown. Light blue circle represents tracking target, yellow circles represent cell number-assigned segments, and green circles represent non-assigned segments. A mitosis was detected at Frame 2.

Supplementary Fig. 1. NIR-DIC and fluorescence light paths, and 2D array.

a. Independent light paths were created for NIR-DIC (Light path 1) and fluorescence (Light path 2) imaging. Light path 1: LED-generated NIR light was passed through a polarizer, Nomarski prism, and condenser to illuminate the sample. The returned light was captured by the CCD camera 1 (512×512 pixels) after passing through an objective lens and Nomarski prism. Light path 2: A laser-generated light was passed through spinning disks (lens and Nipkow disks) and an objective lens to illuminate a sample, and excite fluorophores. Emitted light was captured by the CCD camera 2 (512×512 pixels) after passing through a Nipkow disk. If both NIR-DIC and fluorescence were performed, acquired NIR-DIC and fluorescence images were merged into a single multi-layer TIFF file. **b.** Cells were illuminated using NIR in a multi-well chambered slide. A 2D image acquisition array was used to cover areas of interest. Objective lenses with ×10–×40 magnification can be used.

Supplementary Fig. 2. Setting FOVs and controlling focused image selection.

a. *Map* enabled the display of an interactive map to set positions (part of the map is shown). Gray squares: the objective was moved to an FOV of interest in a well, and preliminary imaging, e.g. 4×4 to 7×7 FOVs, was performed using Metamorph. The xyz positions of each FOV that corresponded to an image acquired by the preliminary imaging were saved on Computer 1 and *Map* software read the positions, displayed as gray squares. Red squares: a 2D image acquisition array created by *Map* software is displayed as red squares. The blue dot represents the objective position. **b.** A flow for the creation of a focused image is shown. Acquired images were saved as multi-layer TIFF files containing unfocused and focused stacks. *Focal Image Selection* was used to convert the files into single-layer TIFF files by either selecting a focal plane, creating all-in-focus images, or creating a stacked image (for fluorescence images). **c**. A multi-layer TIFF file, composed of 20–80 z-planes, contained unfocused and focused planes. *Focal Image Selection* was used to create focused images by selecting the focus plane using Sobel edge extraction and pixel standard deviation or creating all-in-focus images by extracting focused pixels from one plane. **d.** *Focal Image Selection* recorded processing type and/or selected plane number for every time point. Processing types could change during long-term live cell imaging, and were subsequently used when performing post-acquisition processing.

Supplementary Fig. 3. Contrast adjustment, background correction, and image stitching.

a. A flow for the background correction and stitched image creation is shown. Image contrast was adjusted using *Contrast Set*. Customized background patterns were prepared to normalize backgrounds, and stitched images were created after background correction and positional adjustment for each FOV. **b.** Customized backgrounds were prepared using the images recorded in single-layer TIFF files. The 512×512-pixel image was converted into 64×64 blocks, and objects (e.g. indicated by arrows) were then removed by filling pixels with values from pixels right and left, or up and down from the object. Customized backgrounds were created by removing all objects. **c.** *Contrast Set* adjusted contrast, created customized backgrounds, assigned a customized background to relevant FOVs, adjusted FOV

position, and generated stitched images. Processing conditions could change during long-term live cell imaging, and individual recorded conditions were employed for any post-acquisition processing as required. **e.** Stitched image before background correction is shown. **f.** Stitched image after background correction is shown.

Supplementary Fig. 4. Object Tracking Controller.

a. *Object Tracking Controller* main window is shown. This software assigned the cell lineage number, controlled the automatic object tracking process, verified *Automatic Object Tracking* outcomes, and checked and repaired the cell lineage database as required. Lineage and cell numbers entered into the *Queue* list are shown. Images corresponding to the cell were displayed in the *Object Tracking Controller* window by interactively selecting a cell, which corresponded to a segmented area, from the queue. **b**. Interactive cell tracking data verification window is shown. Light blue circle is a segmented area that is being tracked. Green circles are segments with assigned lineage and cell numbers.

Supplementary Fig. 5. Cellular event detection and data analysis.

a. When mitotic cells were identified using interactive drawing tools and by *Automatic Object Tracking*, the cell shapes were stored in the mitotic pattern library. These patterns were used by *Automatic Object Tracking* to detect mitosis by comparing edge circle shapes with the stored patterns. *Automatic Object Tracking*, first detect mitosis and, if two, three, or four tracking targets were found following the mitosis detection, *Automatic Object Tracking* determined if bipolar, tripolar, or tetrapolar cell division had occurred, respectively. Cell death was detected when the bright area persisted across time point images. Cell fusion was identified when two edge circles, each representing a cell, were attached together for a certain period. **b.** *Data Analysis* held multiple cell lineage databases allowing various analyses, including the determination of growth curves and cell doubling times. Default analyses, e.g. drawing cell lineage maps and calculating the numbers of cellular events in each database, were

performed and displayed in the *Data Analysis* window. *Data Analysis* also provided a platform for additional types of analysis as required.

Supplementary Fig. 6. Adjustment of fluorescent image positions.

a. A flow for the adjustment of fluorescent image positions is shown. Fluorescent image positions were adjusted to merge fluorescent images with the live cell imaging video archive. Snapshot 1 was taken before removing the chambered slide from the microscope stage, and the chambered slide was returned to the stage after fixing and staining the cells. Snapshot 2 was then taken to adjust the image position. **b.** *Map* was used to mark the identified identical object positions in Snapshots 1 and 2, calculate corresponding x and y position shifts, and create new FOV positions. **c.** A fluorescent image was overplayed onto NIR-DIC image, which was displayed by *Movie viewer*. In this example, nuclei stained with DAPI were shown.

Supplementary Fig. 7. Automatic object tracking using the computerized single-cell lineage tracking analysis system.

Typical single-cell tracking outcomes at **(a)** time 1 and **(b)** time 693 are shown. Lineage numbers are shown in unique colors. **c.** After staining the cells with TRITC-SNA-I, the sample chamber was returned to the microscope stage, image position was adjusted, and fluorescence images acquired. SNA-I^+ cells were then visualized by fluorescent imaging. **d.** An example of cell lineage maps for a SNA-I^+ lineage is shown. Binding levels of SNA-I are shown at the right end of cell lineages (red, positive; blue, negative; and green, intermediate). Expression levels of parent of SNA-I^+ cells were simulated. For example, when a parent cell produced both SNA-I^+ positive (100%, red line) and SNA-I^- (0%, blue line) cells, the expression level of the parent cell was calculated to be 50% (green line). If a cell showed 50% SNA-I binding and its sibling showed 100%, the SNA-I binding level of the parent cells was 75% (yellow line). The light blue line represents 25% of binding level. Arrows indicate cell

division that produced an SNA-I^+ or SNA-I^- cell. **e.** A simulation was performed based on the assumption that SNA-I^- progeny are produced from a SNA-I^+ progenitor cell. **f.** A simulation was performed based on the assumption that SNA-I^+ progeny are produced from a SNA-I^- progenitor cell.

Supplementary Fig. 8. Cell lineage maps.

Light blue circle, mitosis; orange circle, cell fusion; pink square, cell death; blue square, incomplete cell division; black vertical line, bipolar cell division; red vertical line, multipolar cell division, e.g., tripolar cell division; and orange line, cell fusion. SNA-I staining levels in each cell are shown at the right end of cell lineages (red, positive; blue, negative; and green, intermediate). Horizontal scale: ×10 min.

Supplementary Fig. 9. Cell lineage maps for SNA-I^+ cell lineages.

Cell lineages, any of which progeny found at Time point 693 were SNA-I^+ cells, are listed. Light blue circle, mitosis; orange circle, cell fusion; pink square, cell death; blue square, incomplete cell division; black vertical line, bipolar cell division; red vertical line, multipolar cell division, e.g., tripolar cell division; and orange line, cell fusion. SNA-I staining levels in each cell are shown at the right end of cell lineages (red, positive; blue, negative; and green, intermediate). Horizontal scale: ×10 min.

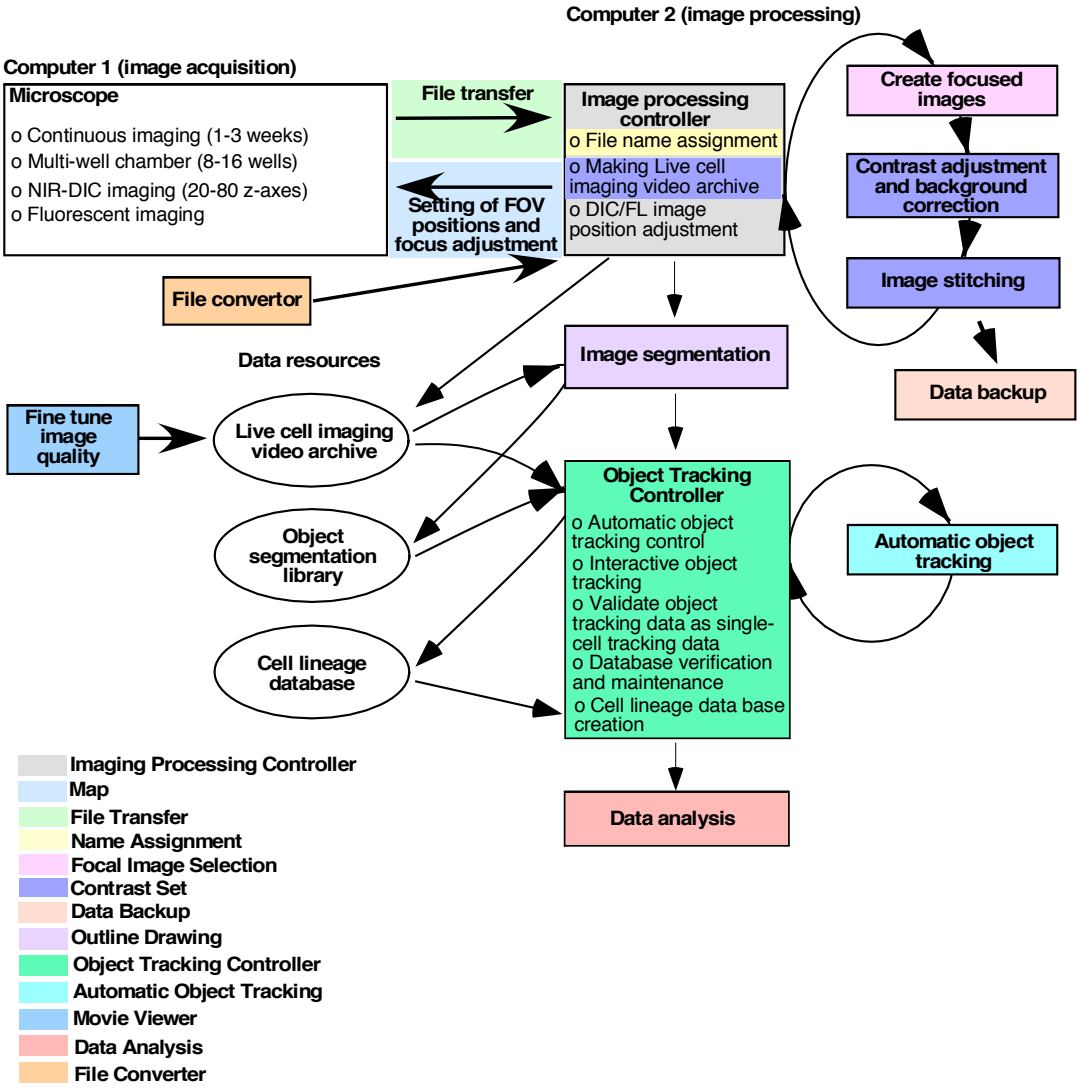
Supplementary Table 1. Developed software (C/C++/Objective-C).

Software	**Functions**
Image Processing Controller	• Control other associated software • Control automatic live cell video making • Merging fluorescent data into the live cell imaging video archive • Control post-live cell imaging processing

File Transfer	• Transfer image files from Computer 1 to Computer 2
Map	• Display well and objective positions • Set two-dimensional image acquisition arrays • Set focal plane z-position
Name Assignment	• Assign a unique file name to each image file
Focal Image Selection	• Create focused images
Contrast Set	• Create customized background patterns • Adjust background • Adjust positions for each FOV • Perform image stitching • Create live cell imaging video archive
Data Backup	• Backup data files
File Converter	• Import imaging files created by other microscope systems • Convert file into suitable format for processing on the system
Outline Drawing	• Perform image data segmentation • Create object segmentation library
Movie Viewer	• Display videos • Fine-tune image position and contrast
Object Tracking Controller	• Manage object tracking process • Control automatic object tracking • Support verification and corrections of object tracking data and event identification using tools, including interactive line and circle drawing tools • Validate object tracking data as single-cell tracking data • Create cell lineage database

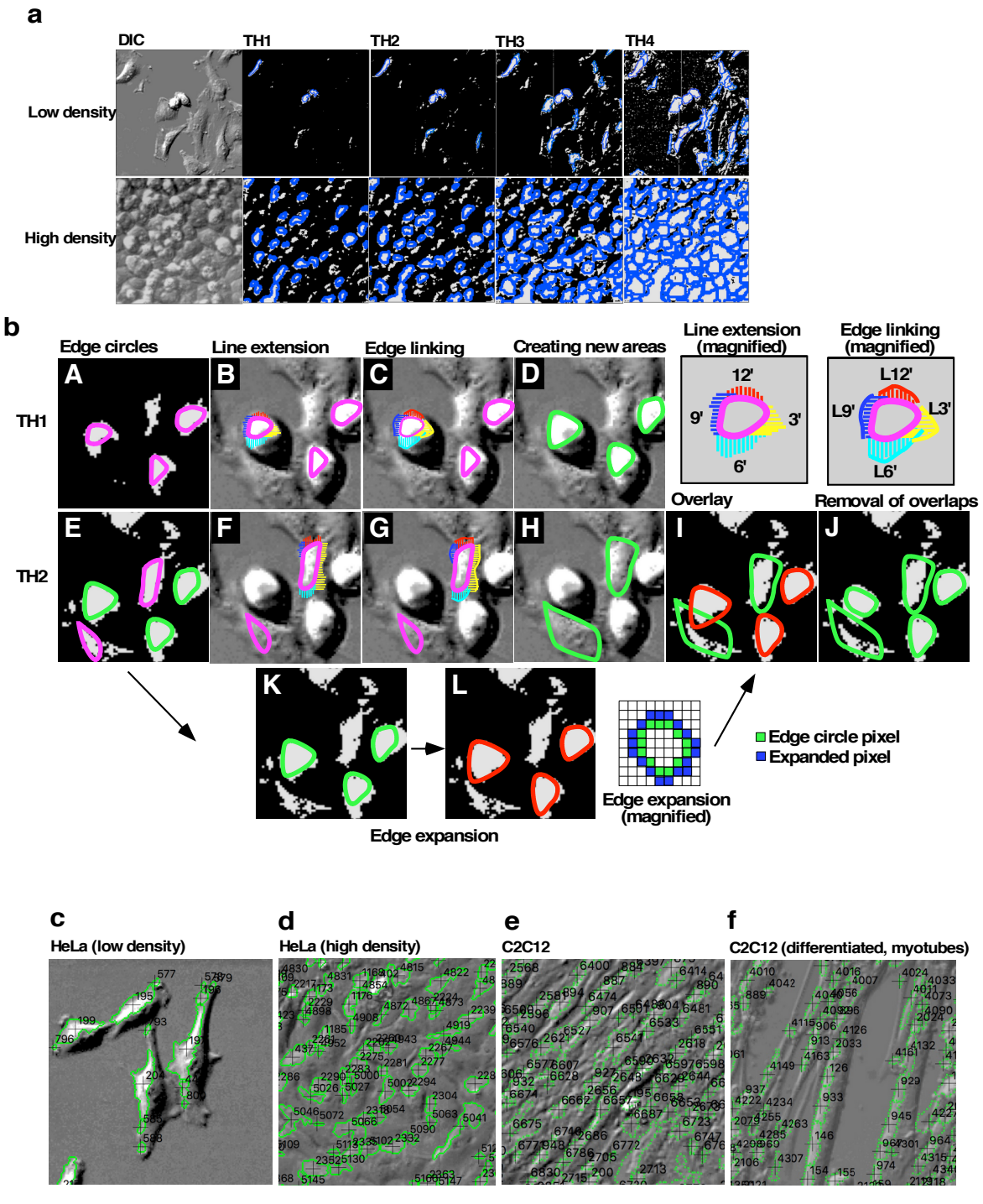
	• Create mitosis pattern library • Verify and repair cell lineage database
Automatic Object Tracking	• Perform automatic object tracking and cellular event detection
Data Analysis	• Perform various data analyses

Fig. 1



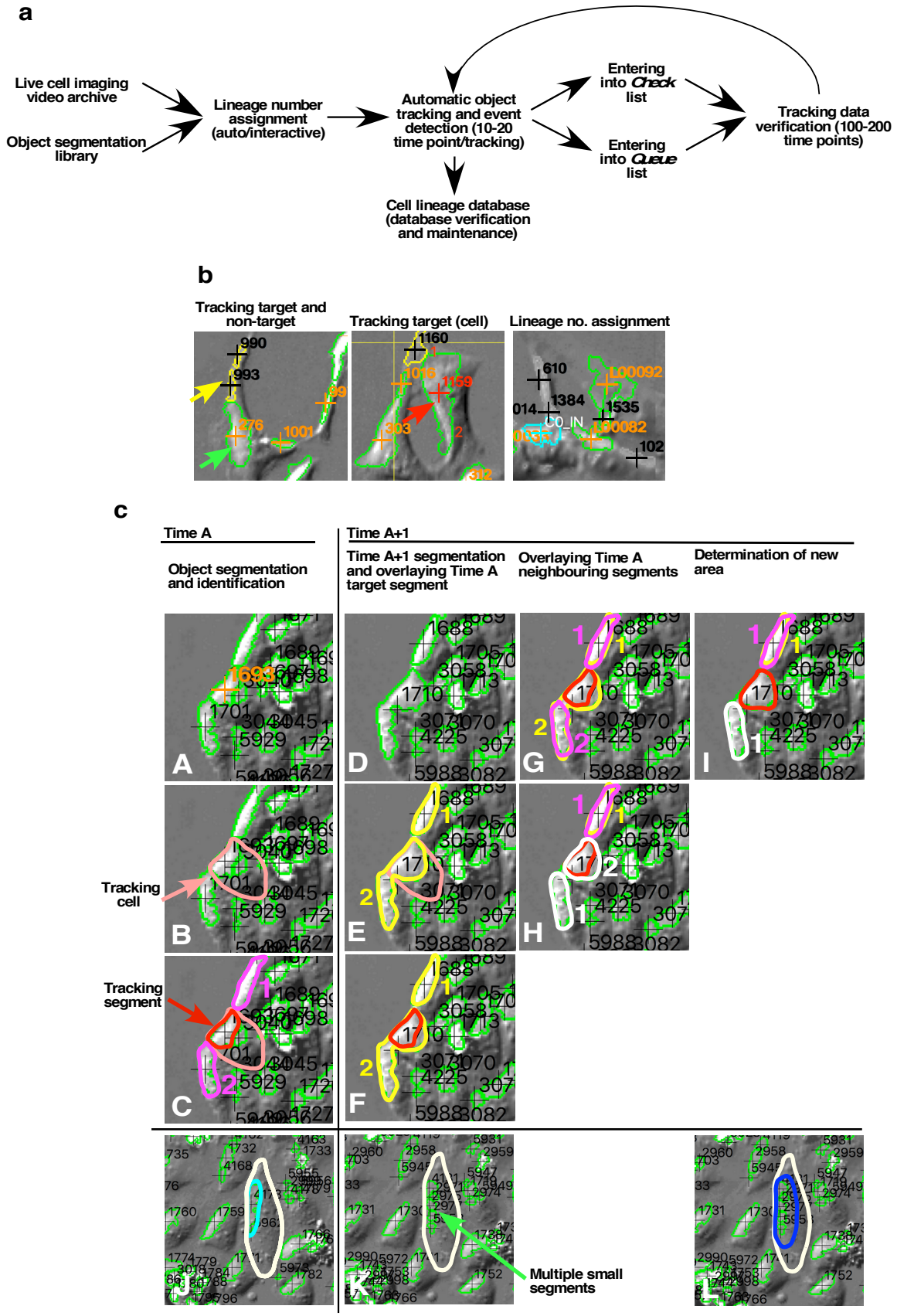
Sato et al.

Fig. 2



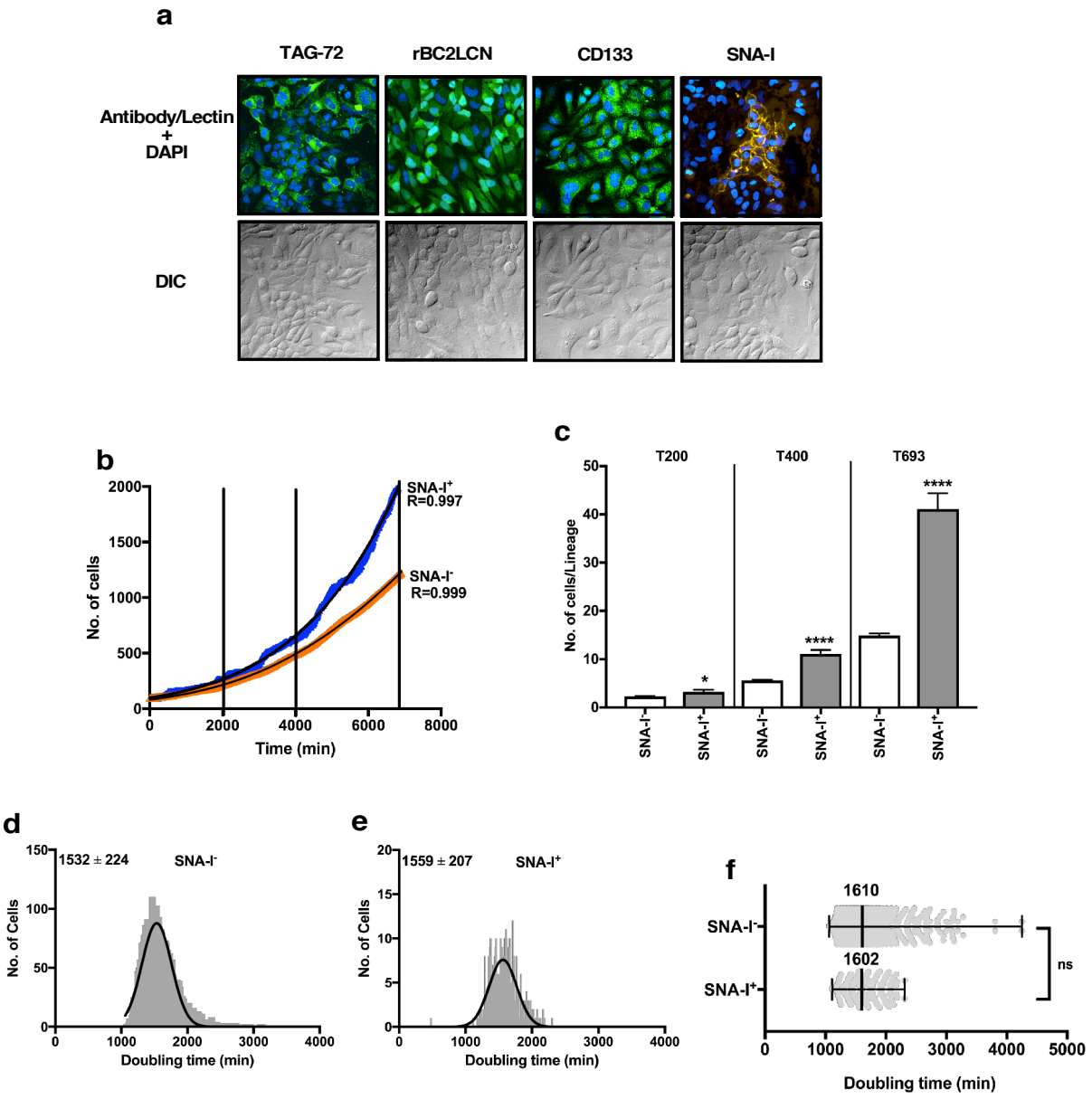
Sato et al.

Fig. 3



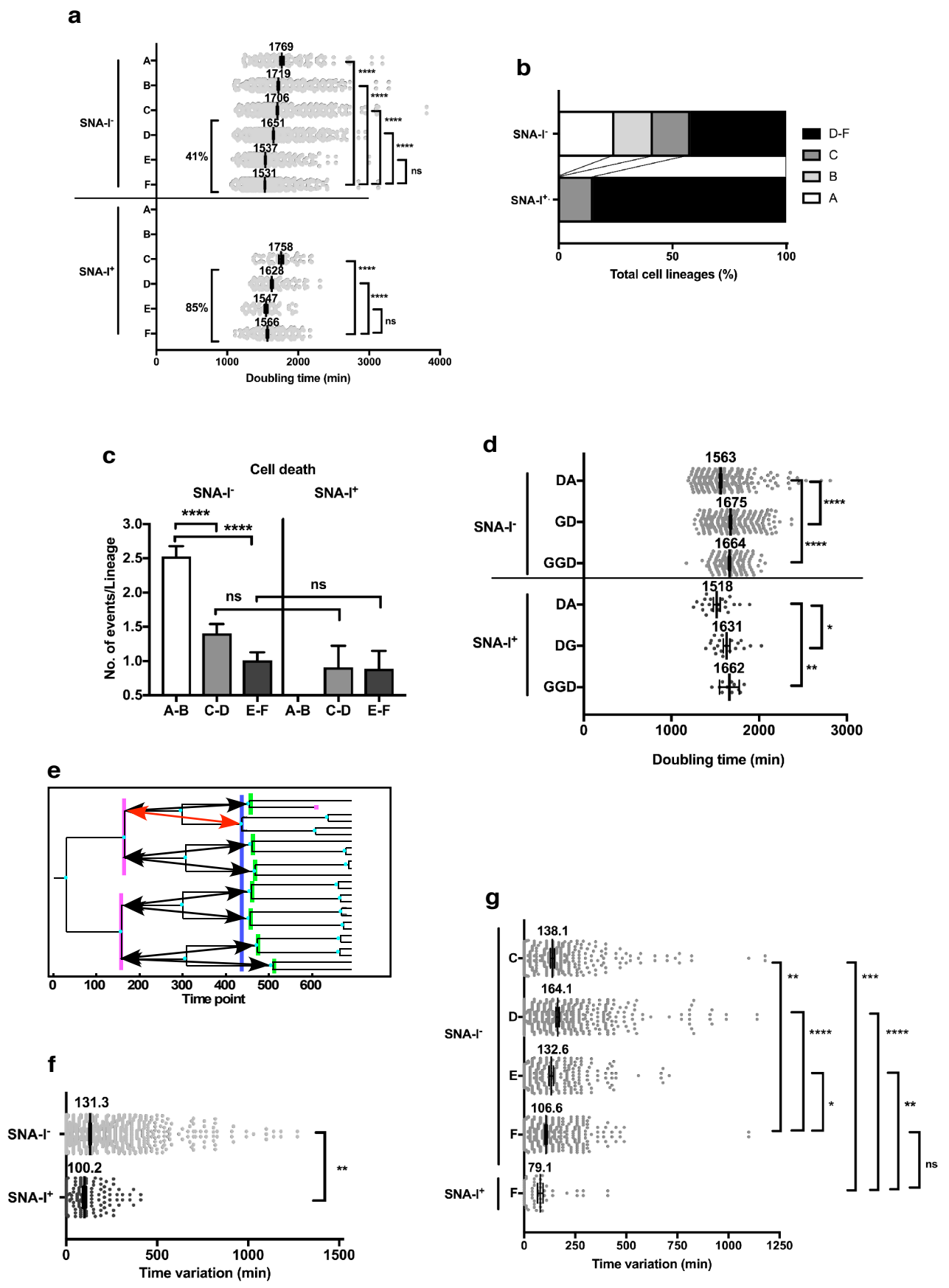
Sato et al.

Fig. 4



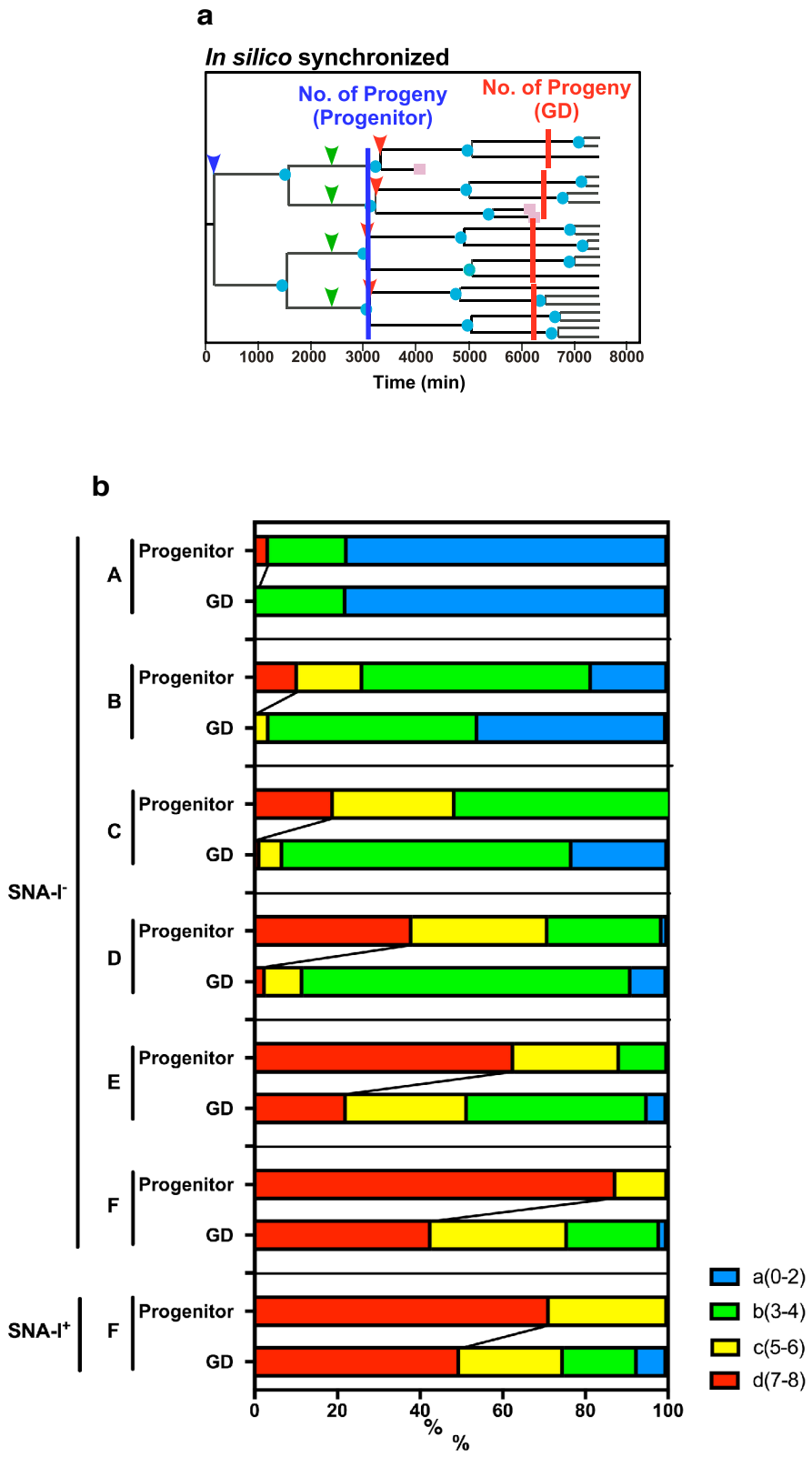
Sato et al.

Fig. 5



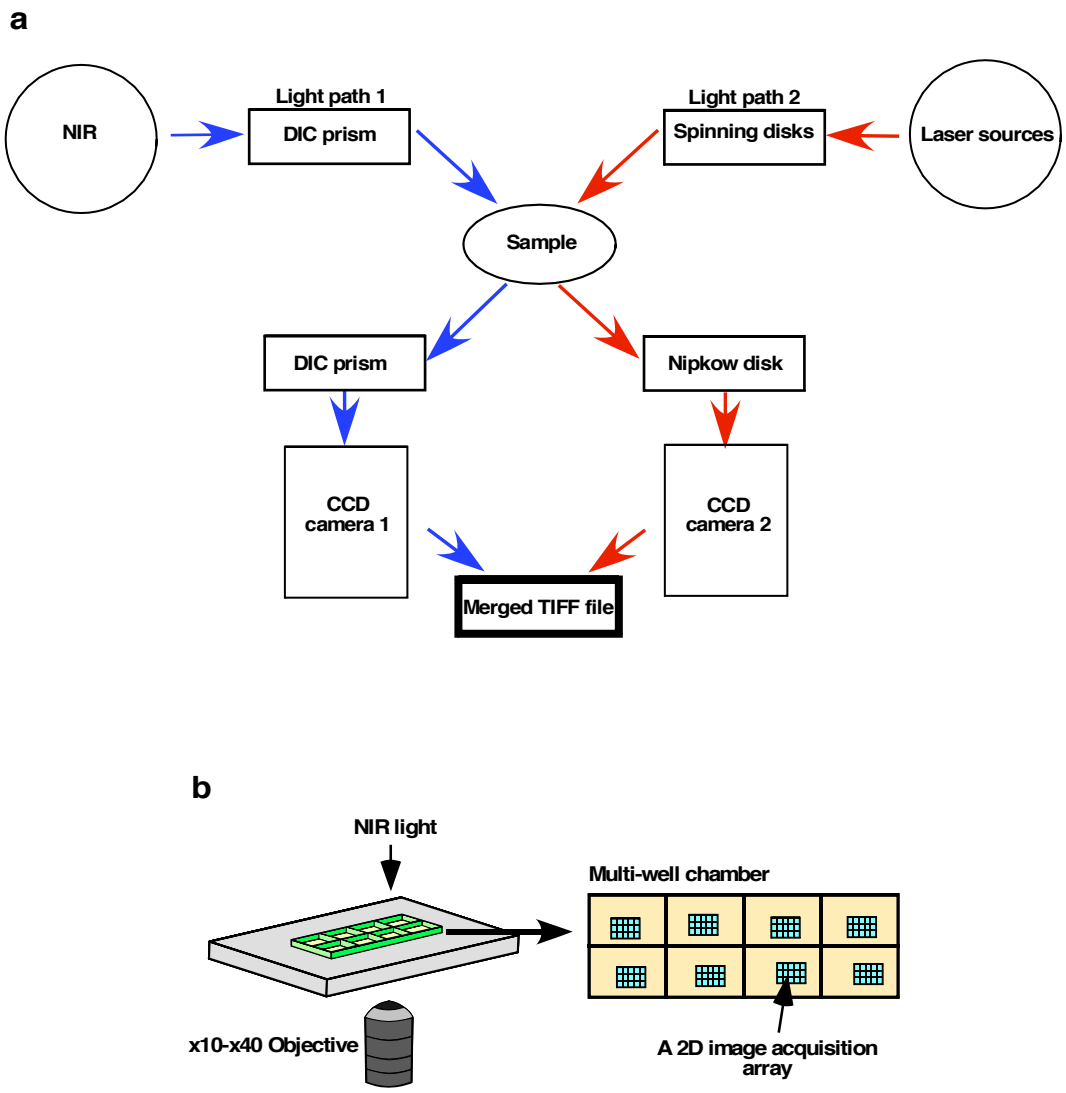
Sato et al.

Fig. 6



Sato et al.

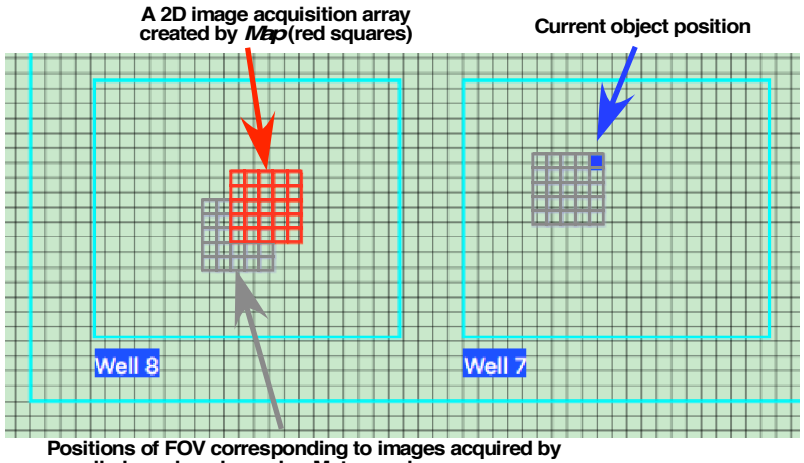
Sup. Fig. 1



Sato et al.

Sup. Fig. 2

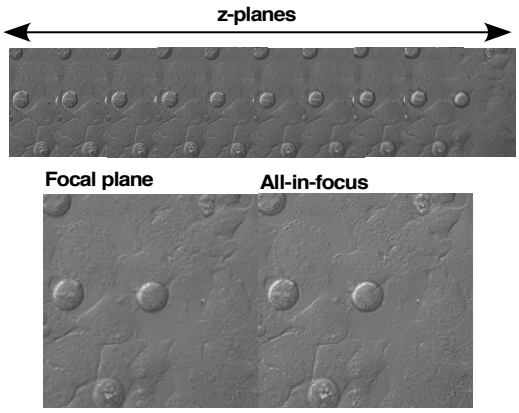
a



Positions of FOV corresponding to images acquired by a preliminary imaging using Metamorph

b

c

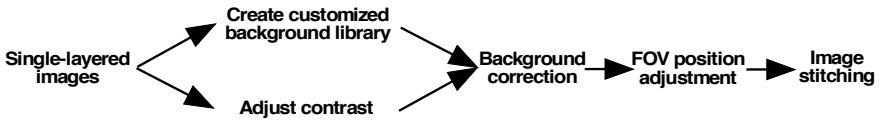


d

Sato et al.

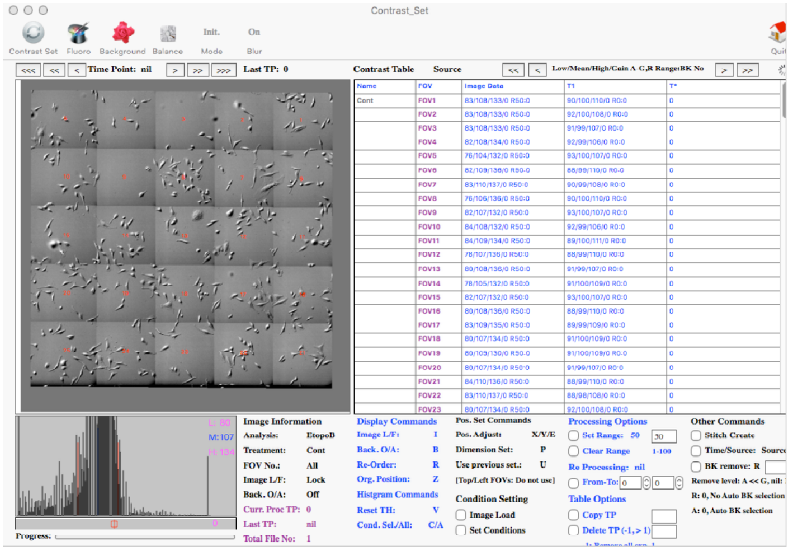
Sup. Fig. 3

a



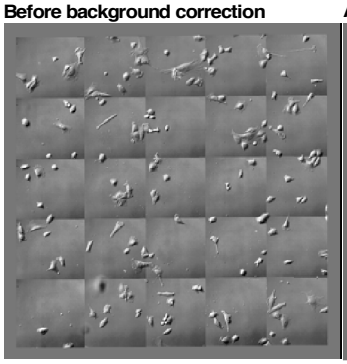
b

c



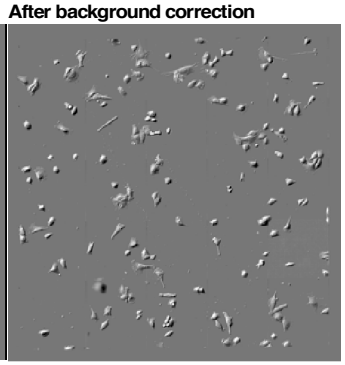
d

Before background correction



e

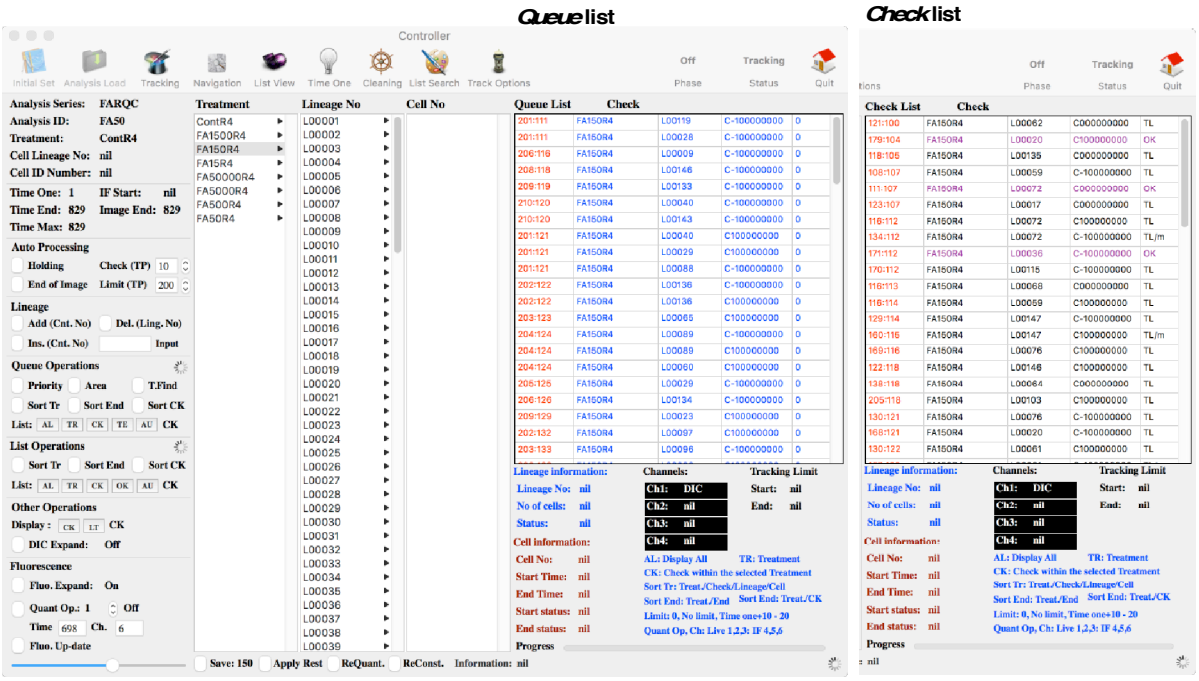
After background correction



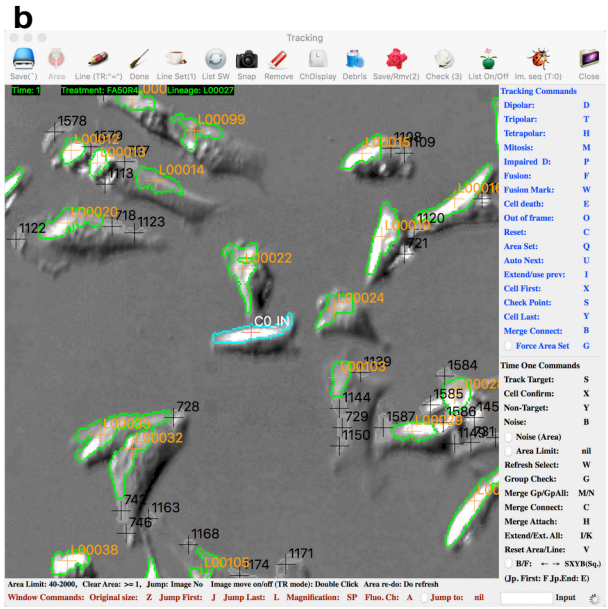
Sato et al.

Sup. Fig. 4

a



b



Sato et al.

Sup. Fig. 5

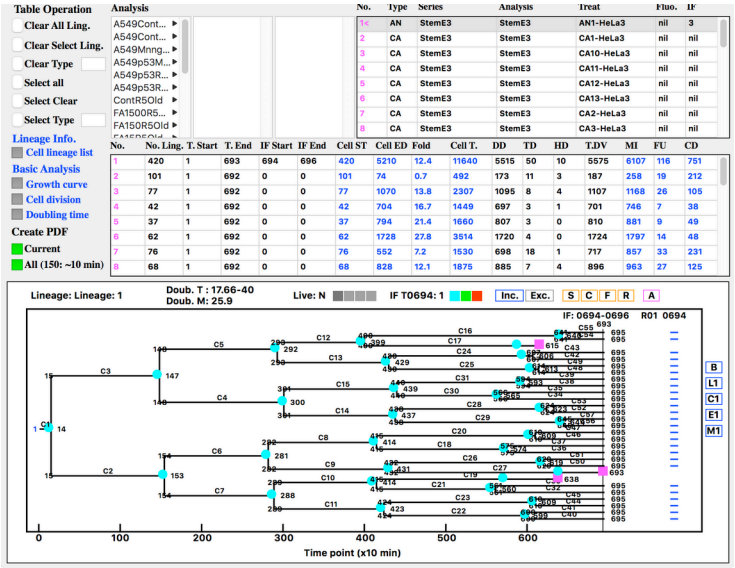
a

Mitosis pattern

Bipolar cell division	Mitosis detection	→	Searching two bright objects
Tripolar cell division	Mitosis detection	→	Searching three bright objects
Tetrapolar cell division	Mitosis detection	→	Searching four bright objects
Cell death	Finding a bright object (non-mitotic)	→	Counting how may time points the object remain
Cell fusion	Checking attached object	→	Setting fusion mark

b

Cell lineage database → Analysis platform → o Cell growth curve
o Cell doubling time
o Event analysis
⋮

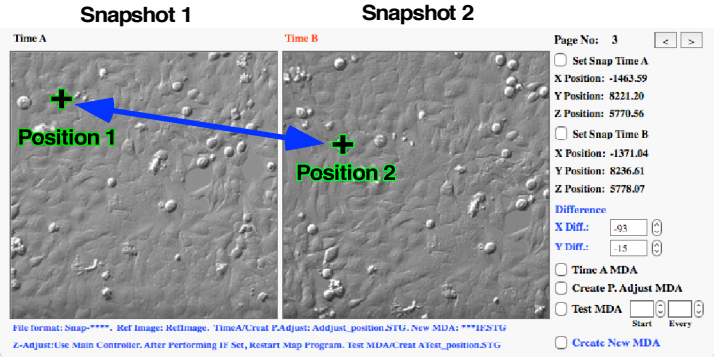


Sato et al.

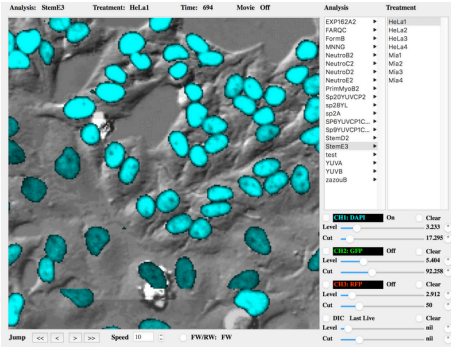
Sup. Fig. 6

a

b



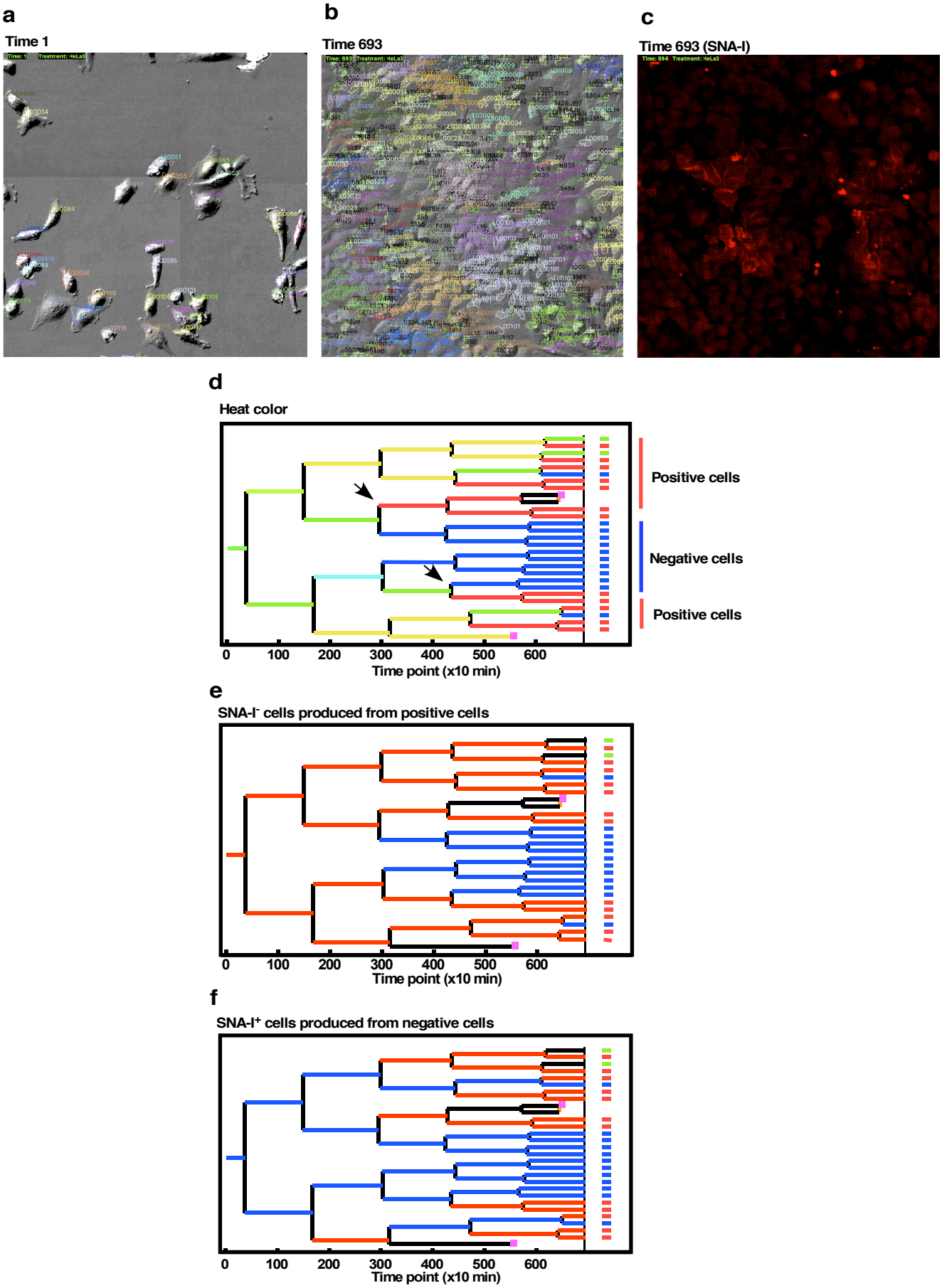
c



Light blue: DAPI-stained nuclei

Satoh et al.

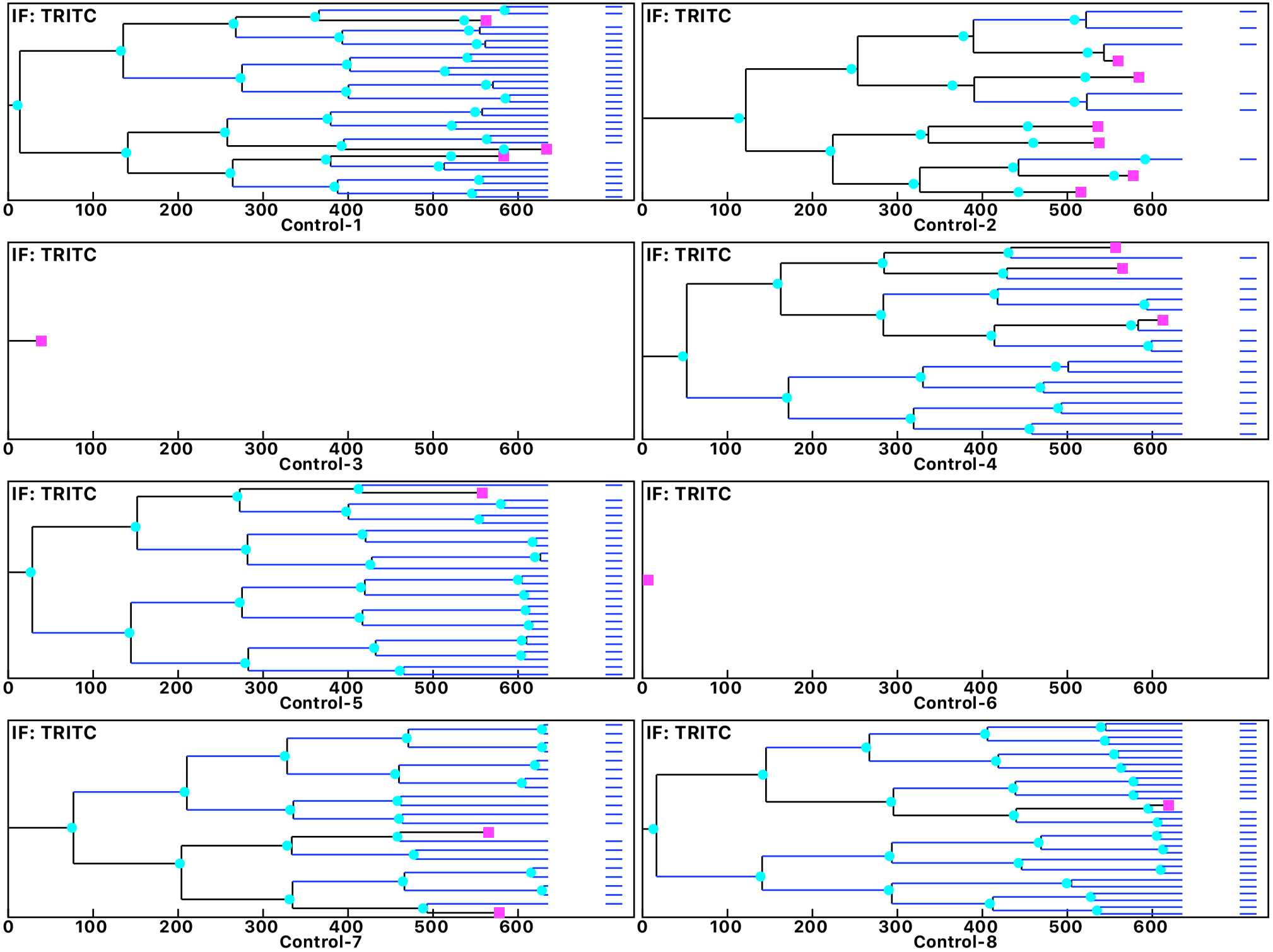
Sup. Fig. 7



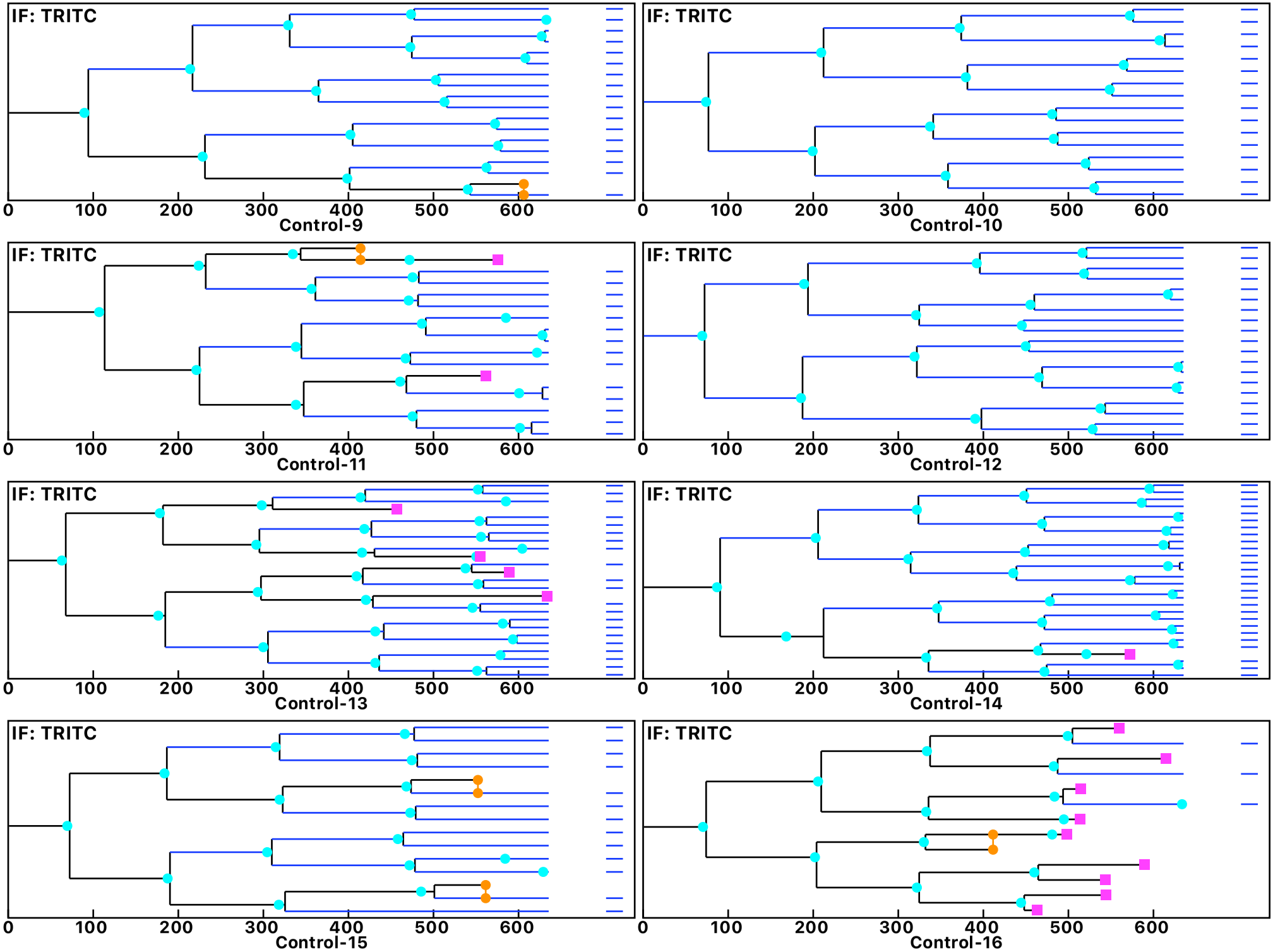
Sato et al.

Supplementary Figure S8

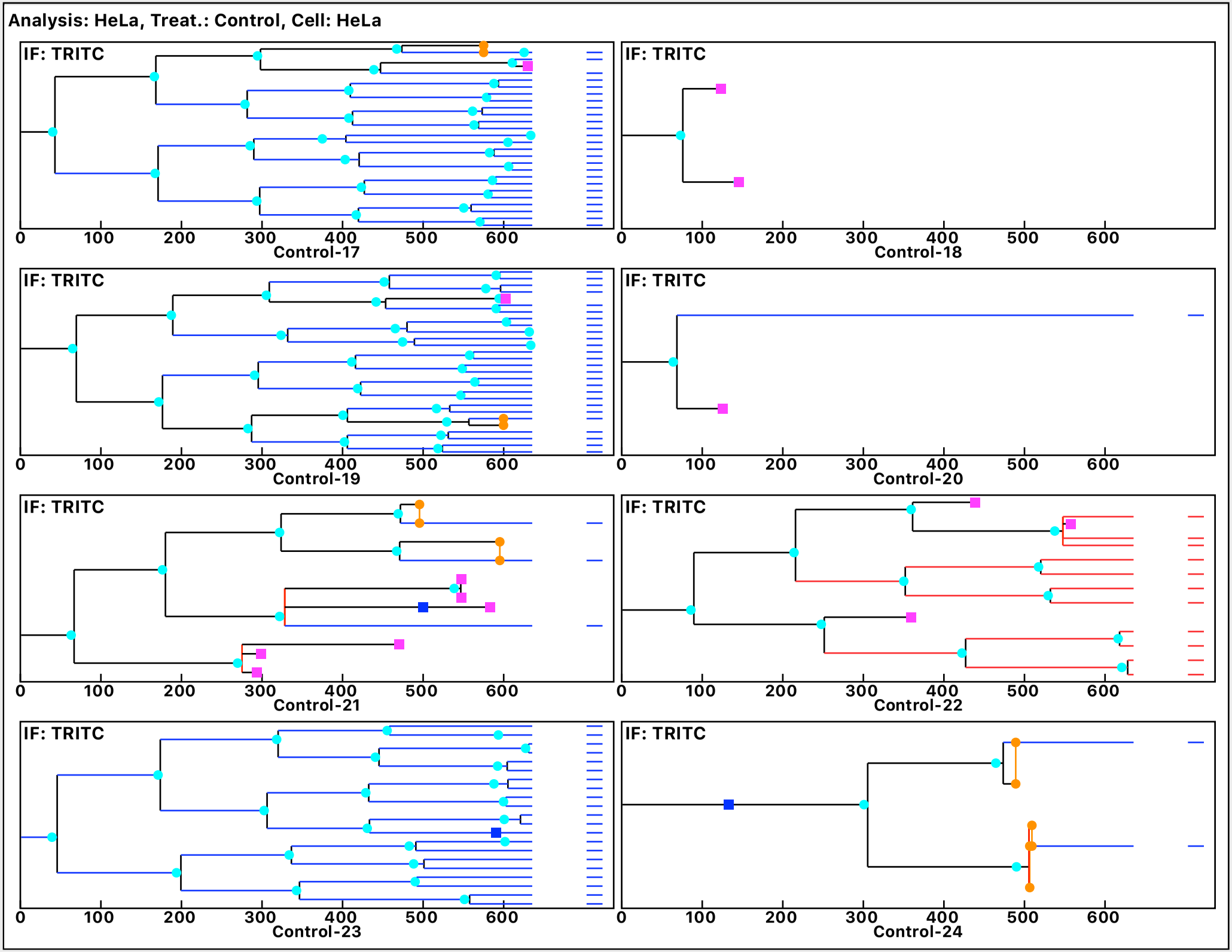
Analysis: HeLa, Treat.: Control, Cell: HeLa



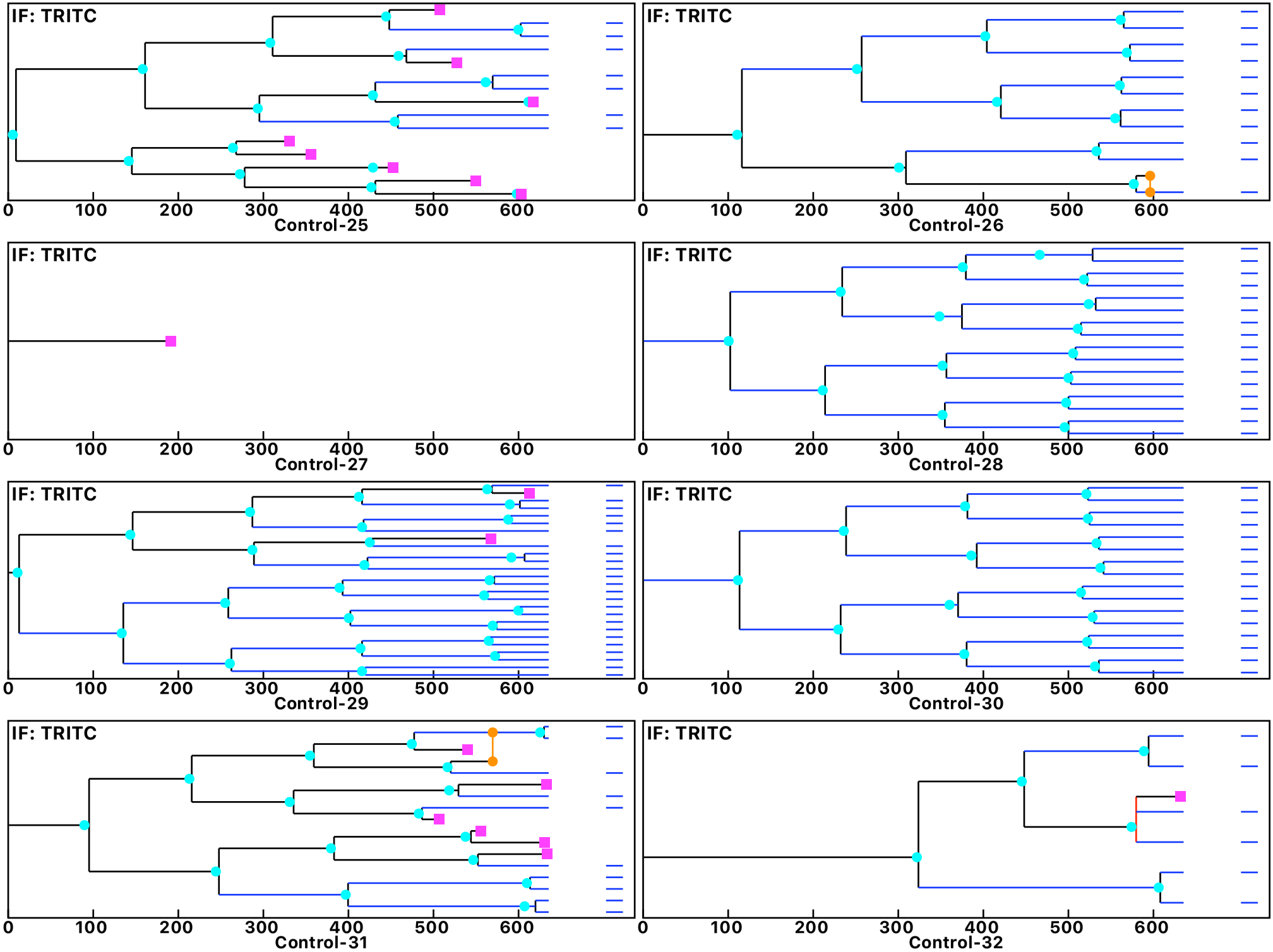
Analysis: HeLa, Treat.: Control, Cell: HeLa



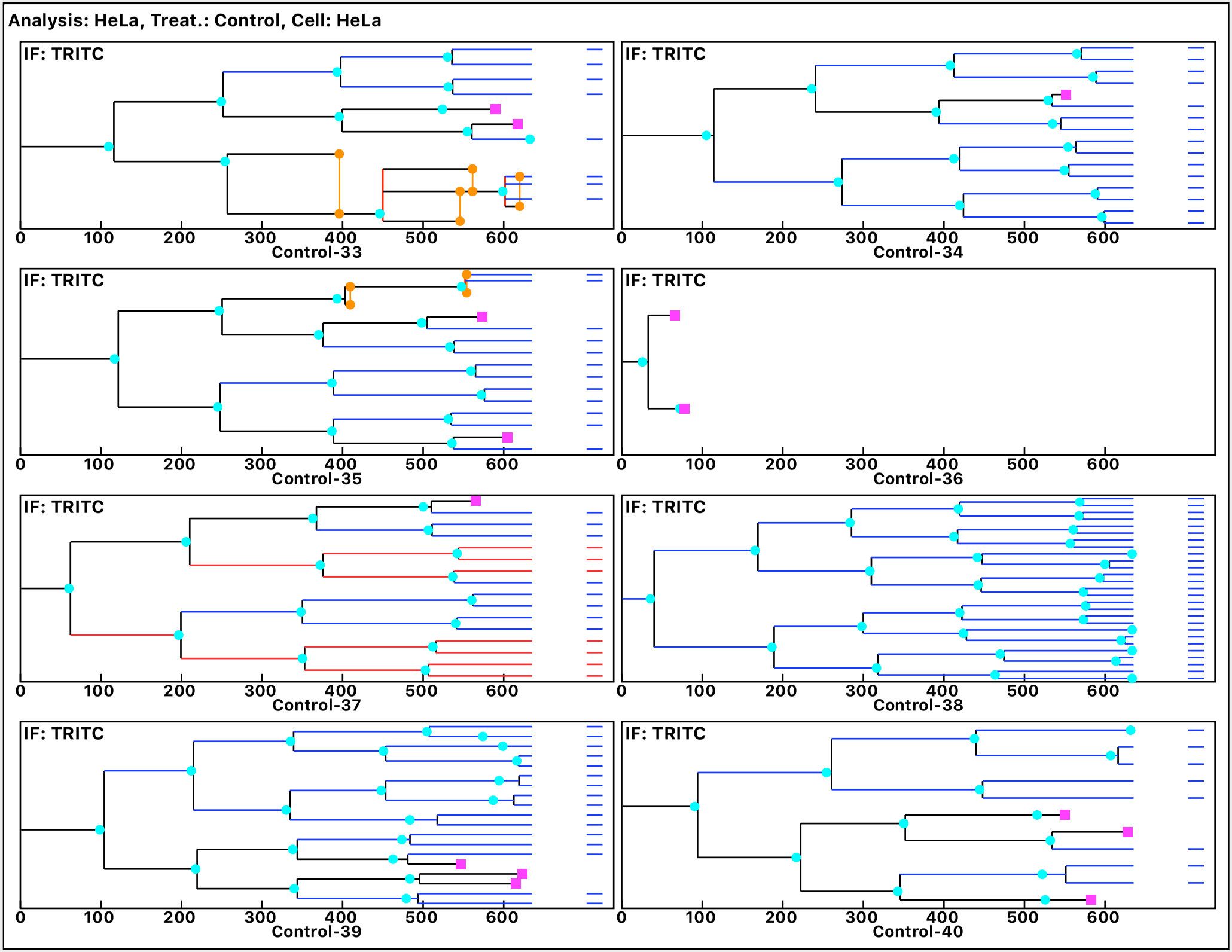
Analysis: HeLa, Treat.: Control, Cell: HeLa



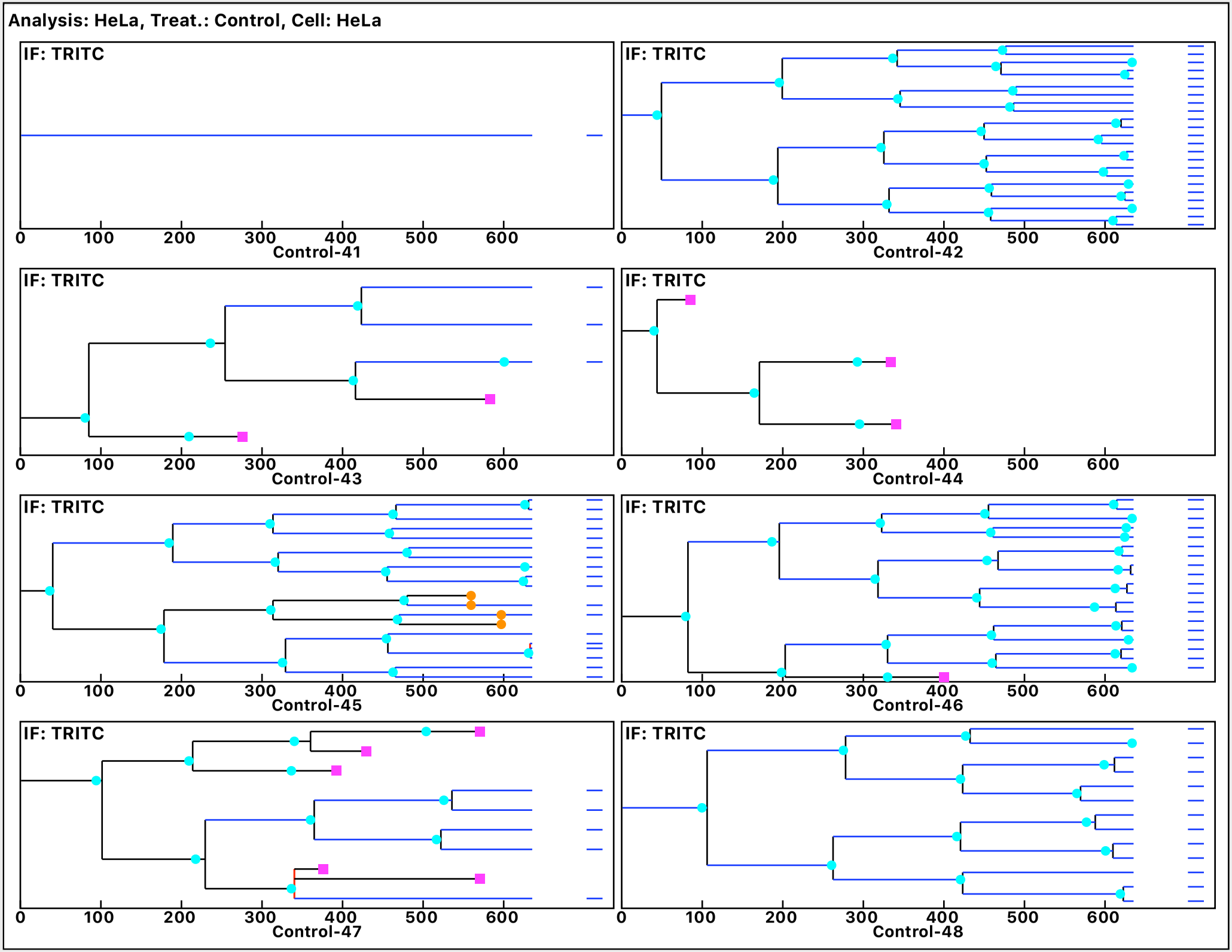
Analysis: HeLa, Treat.: Control, Cell: HeLa



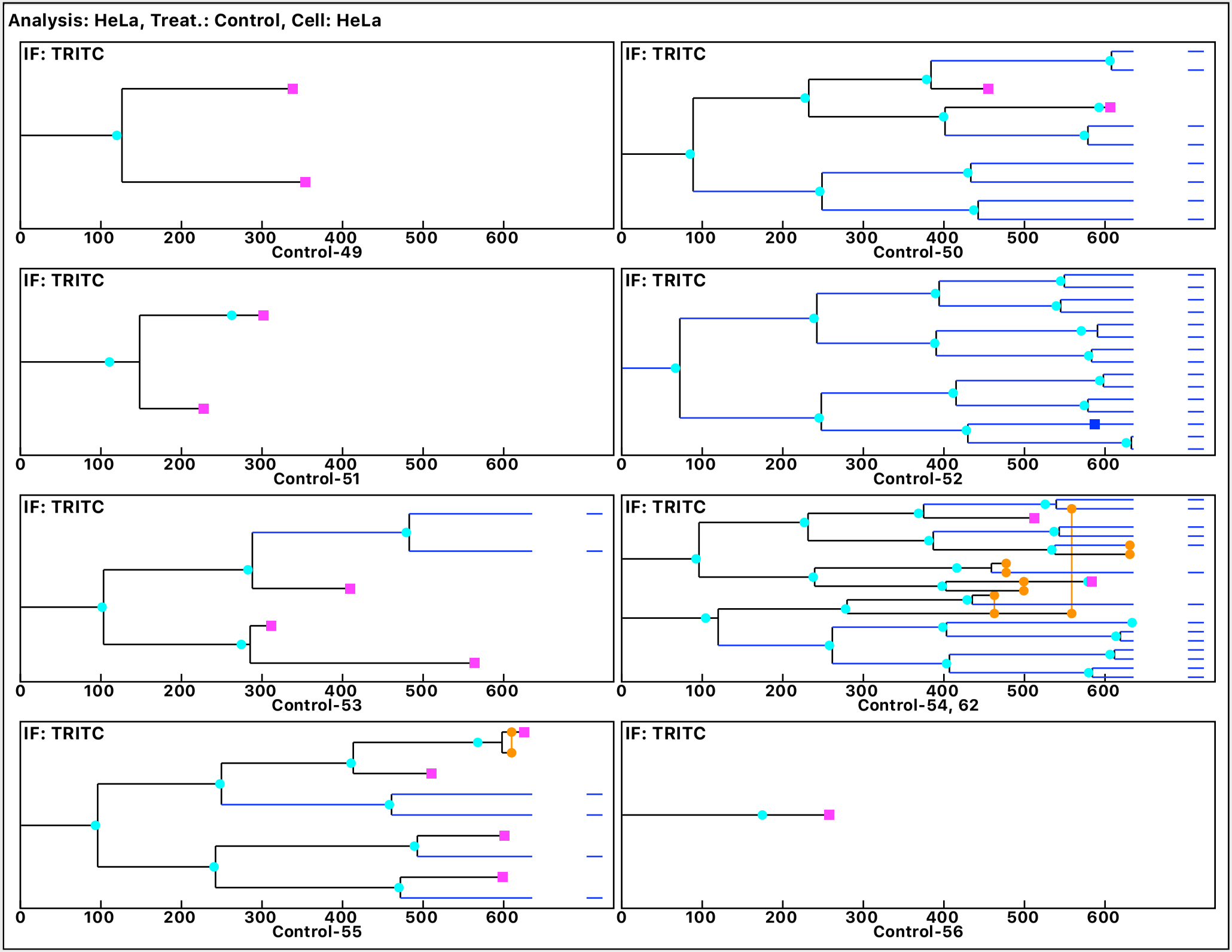
Analysis: HeLa, Treat.: Control, Cell: HeLa



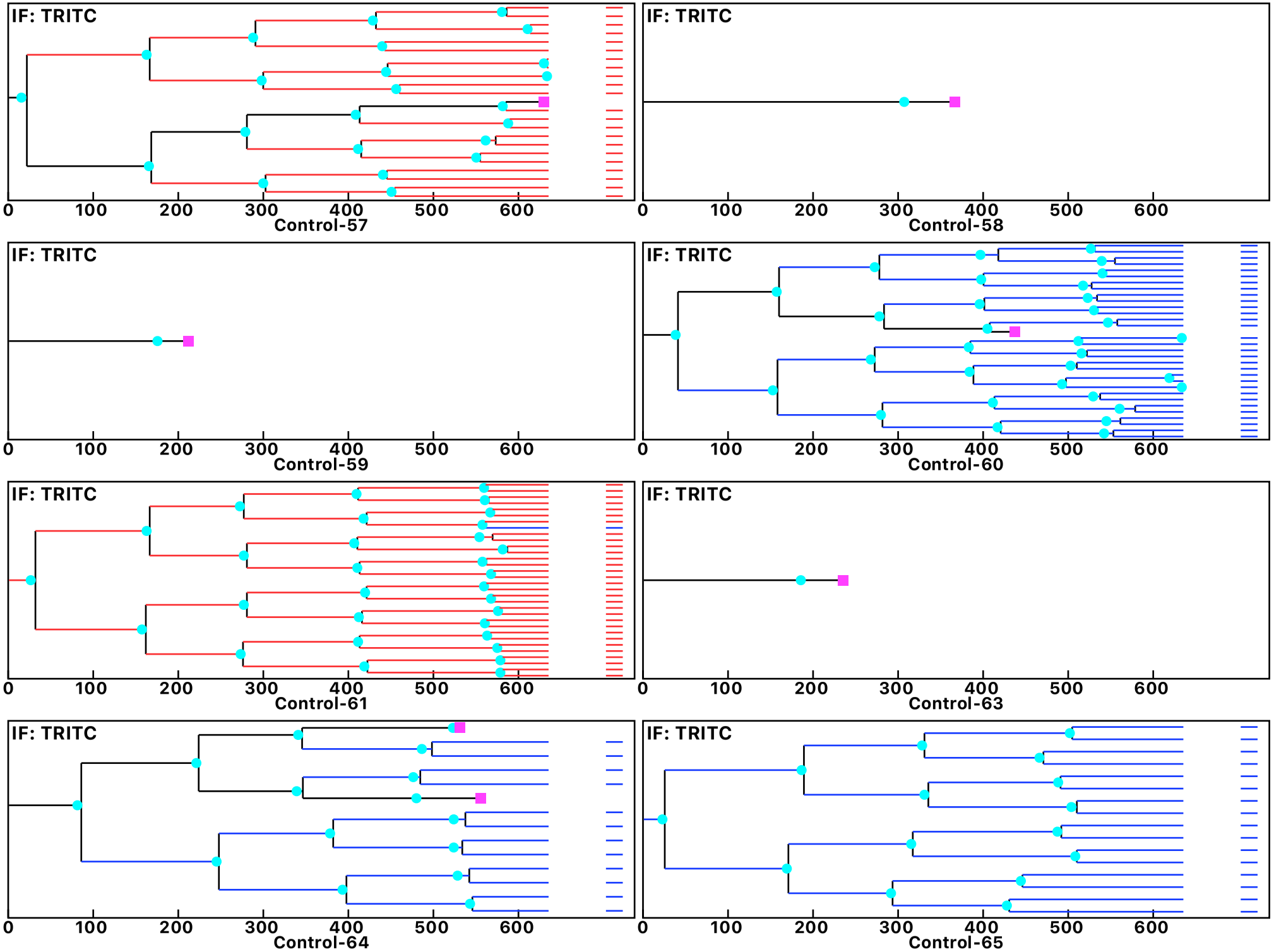
Analysis: HeLa, Treat.: Control, Cell: HeLa



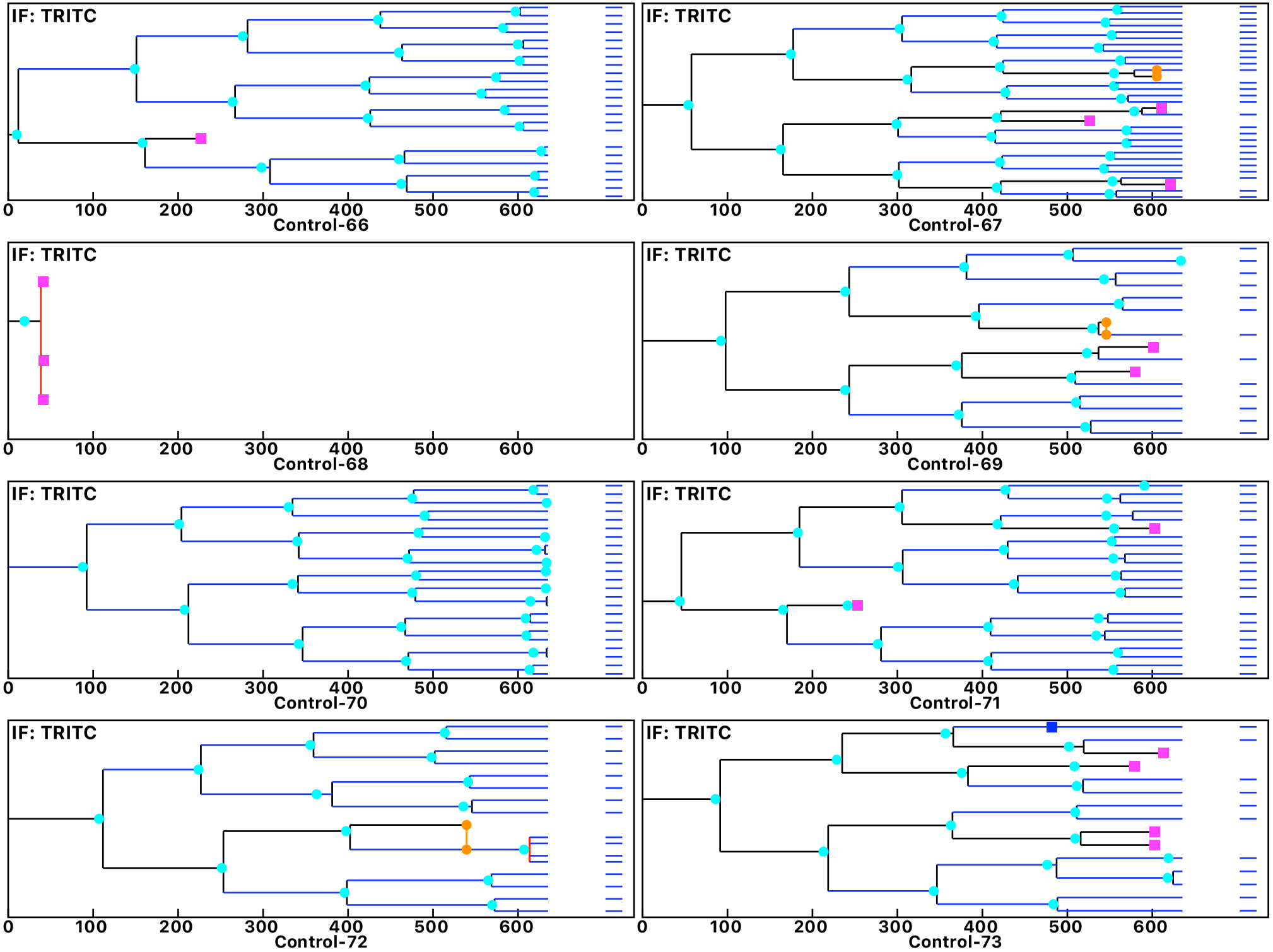
Analysis: HeLa, Treat.: Control, Cell: HeLa



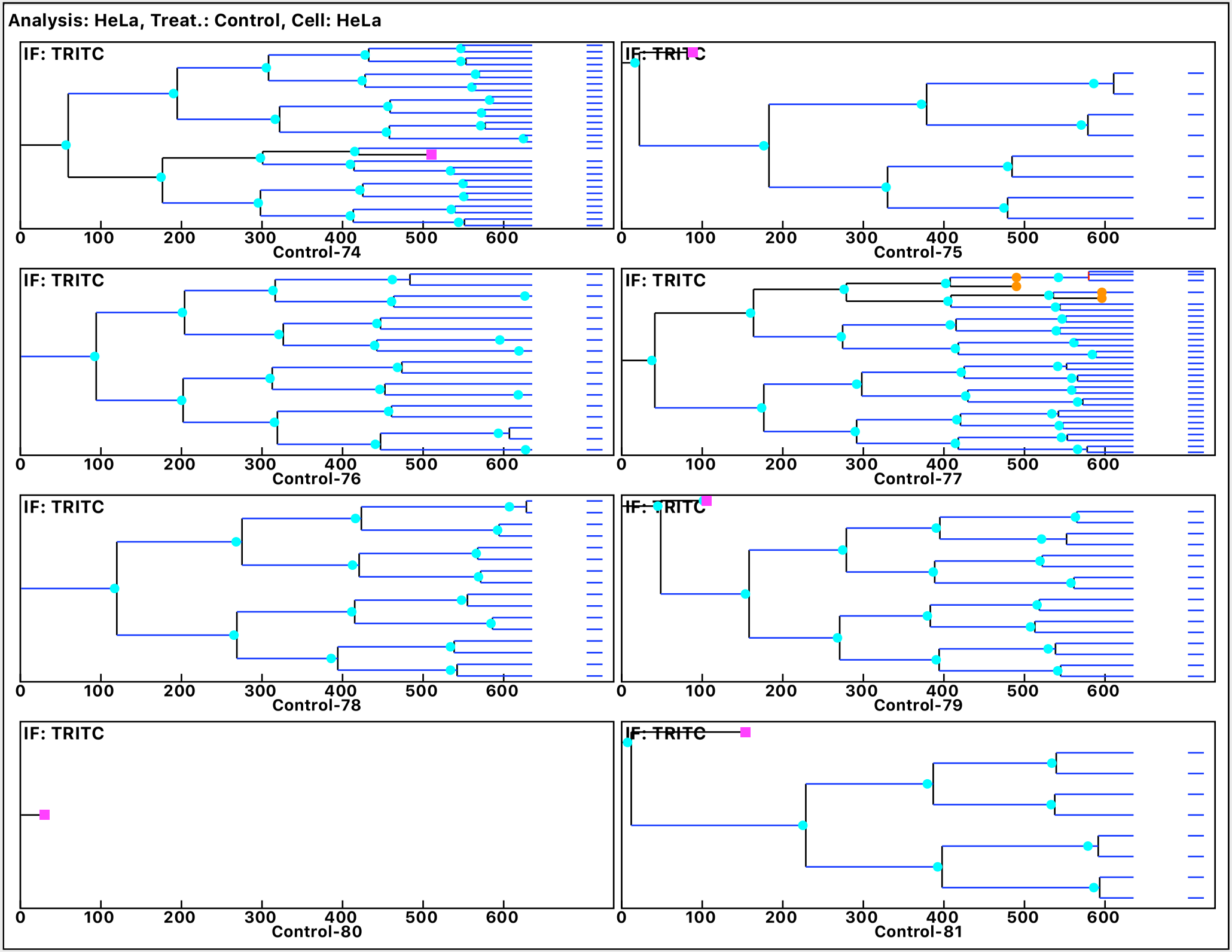
Analysis: HeLa, Treat.: Control, Cell: HeLa



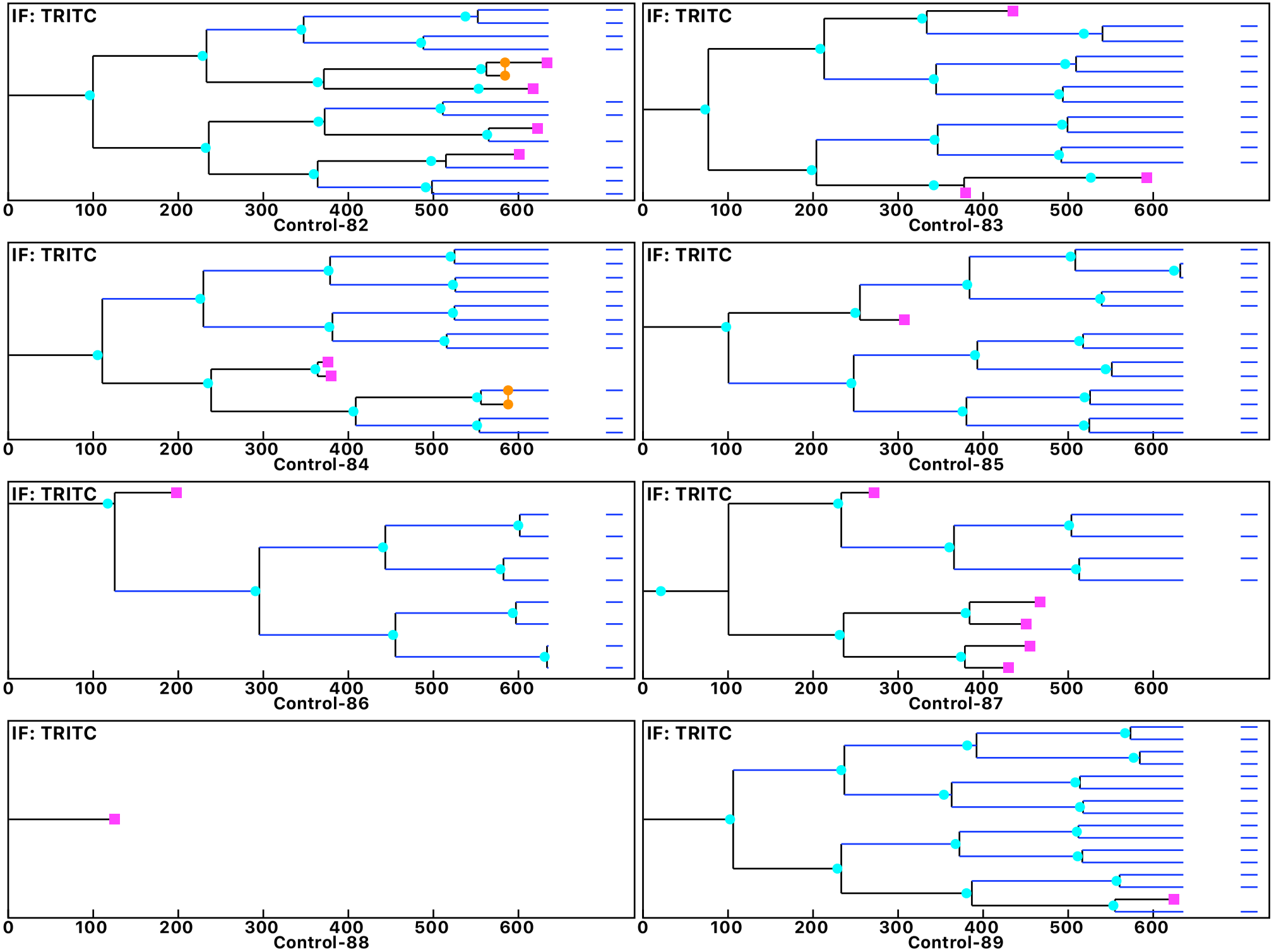
Analysis: HeLa, Treat.: Control, Cell: HeLa



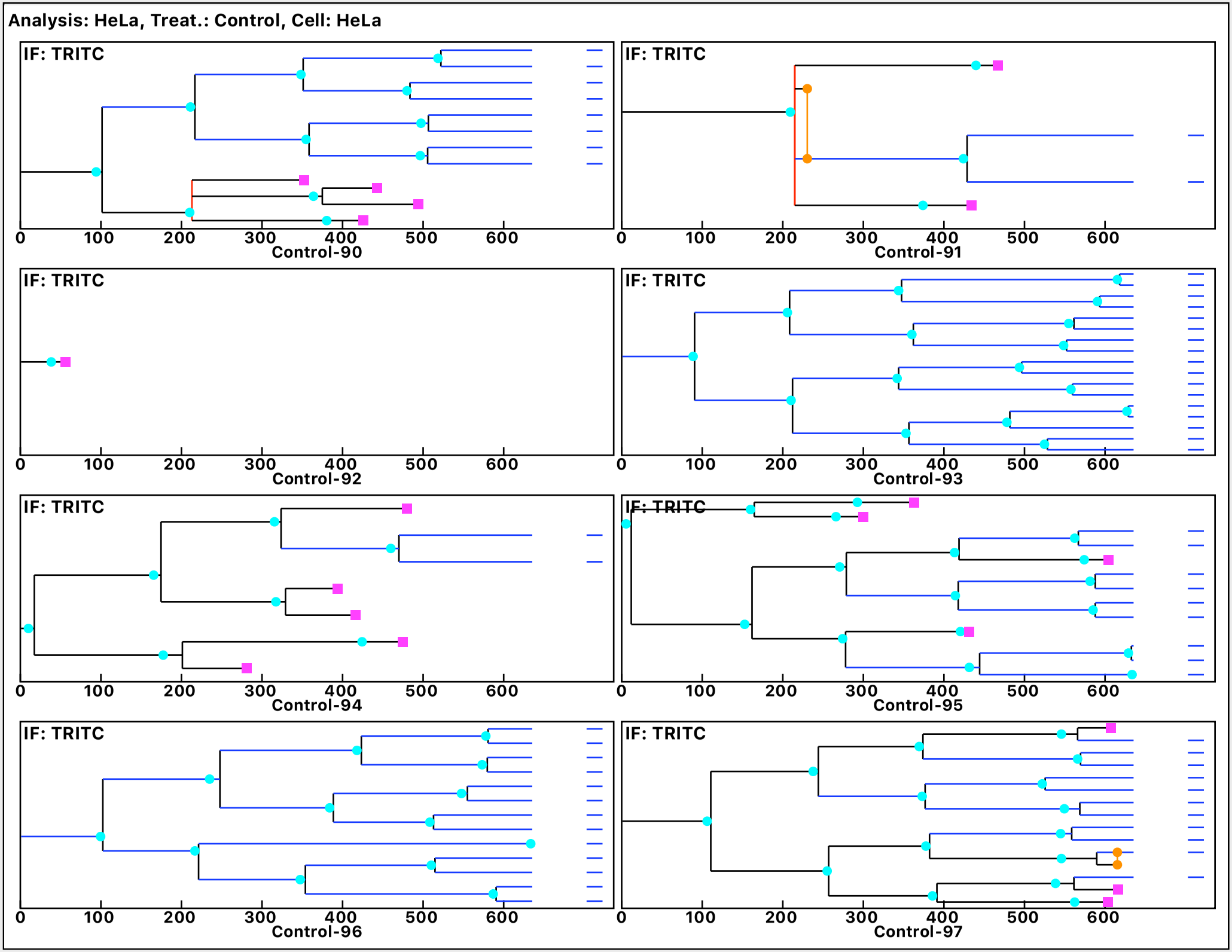
Analysis: HeLa, Treat.: Control, Cell: HeLa



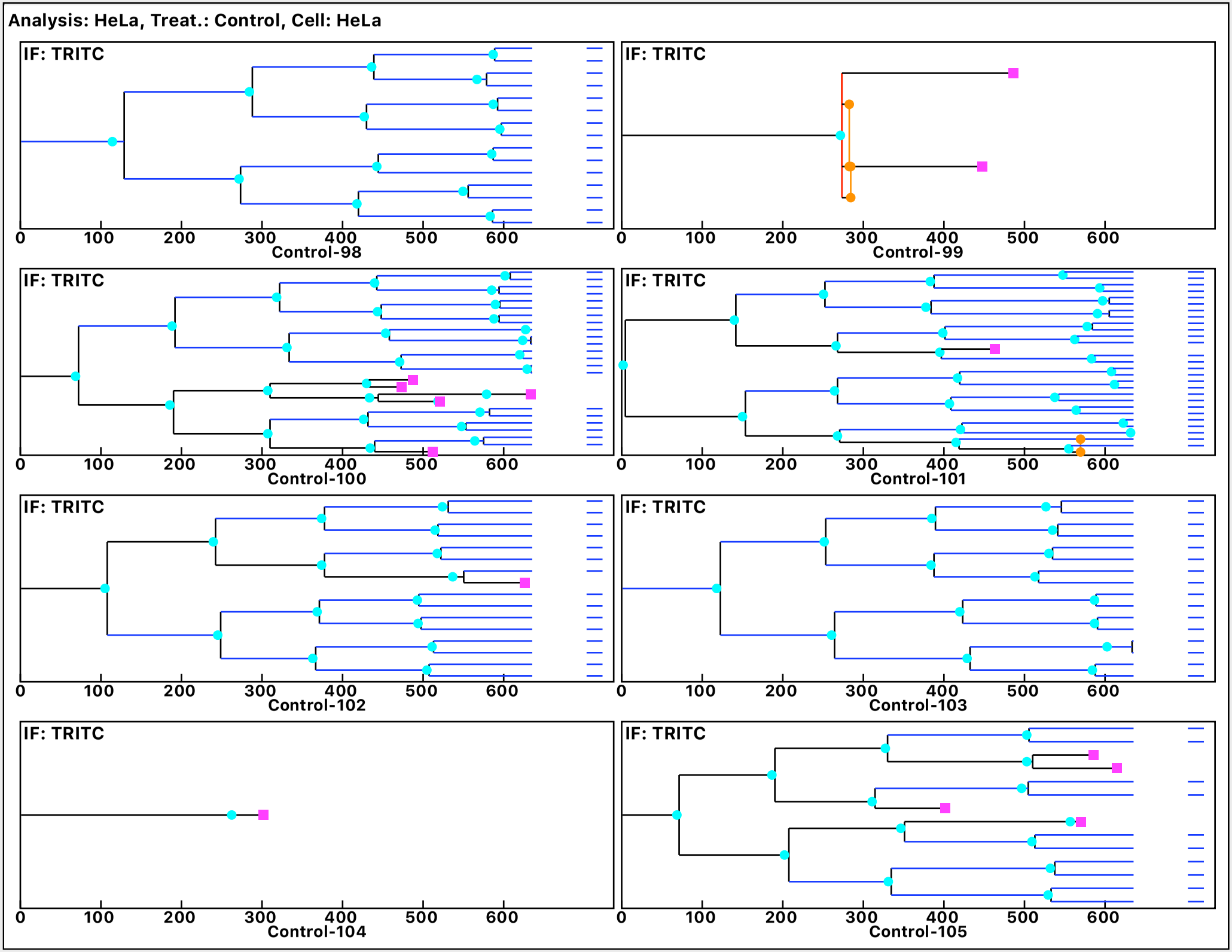
Analysis: HeLa, Treat.: Control, Cell: HeLa



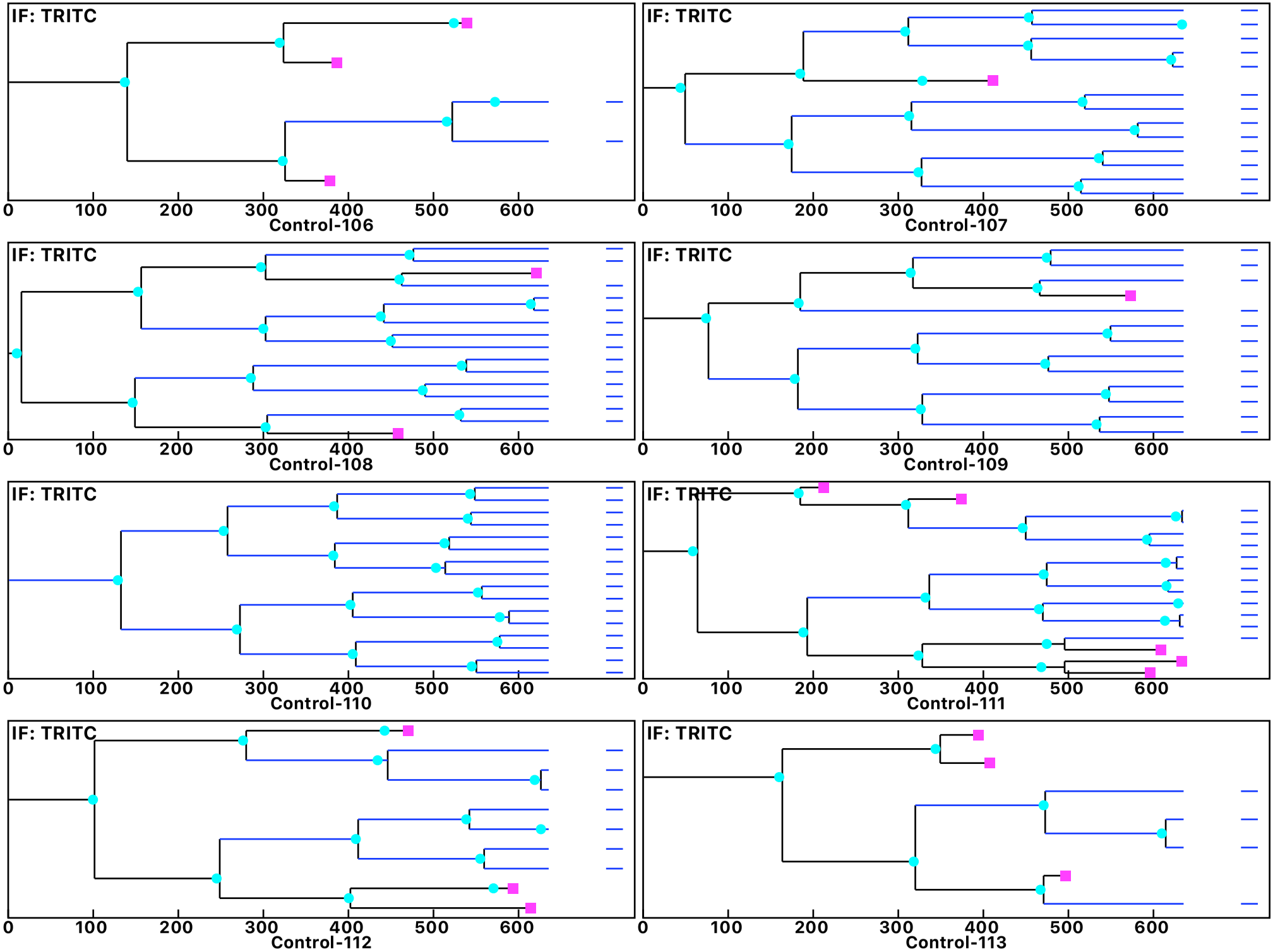
Analysis: HeLa, Treat.: Control, Cell: HeLa



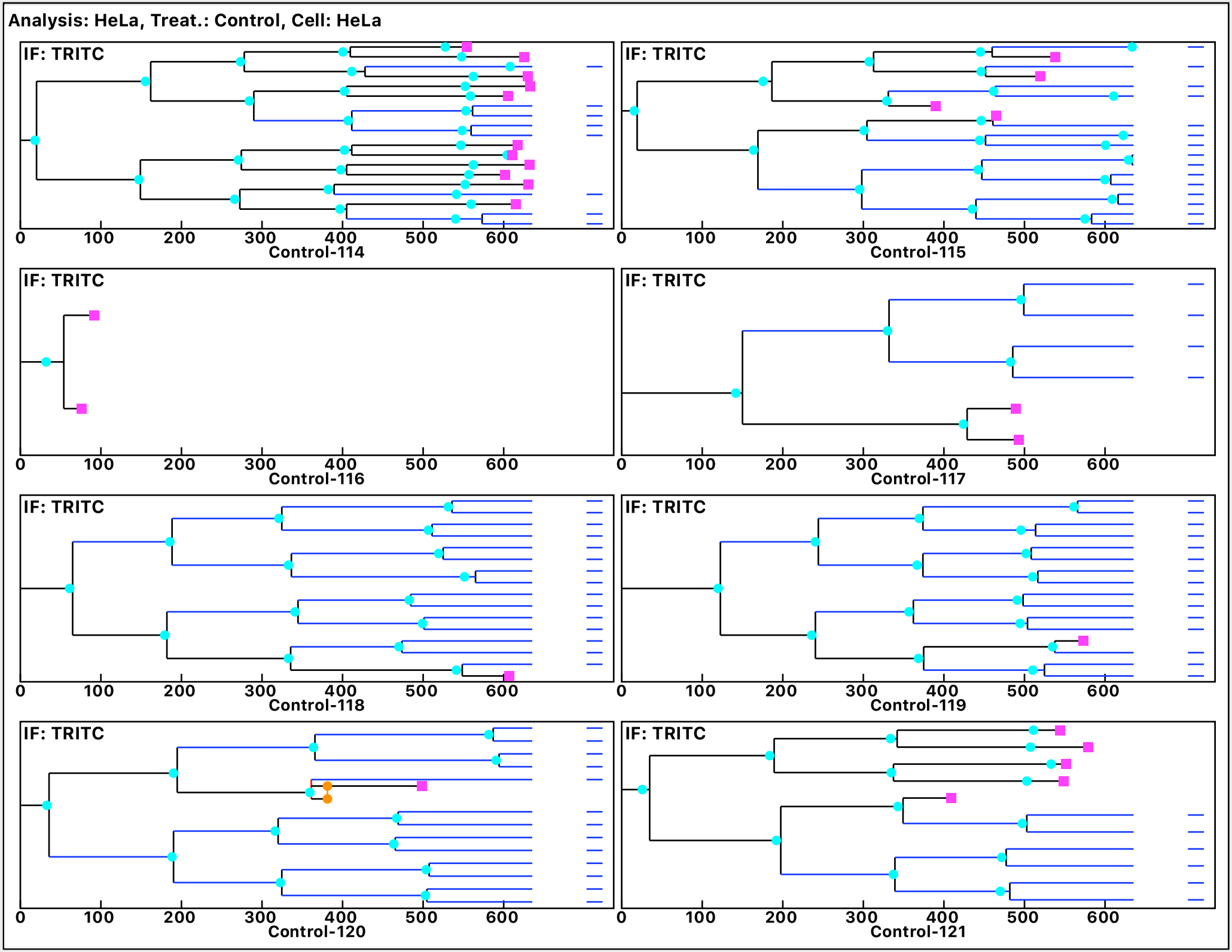
Analysis: HeLa, Treat.: Control, Cell: HeLa



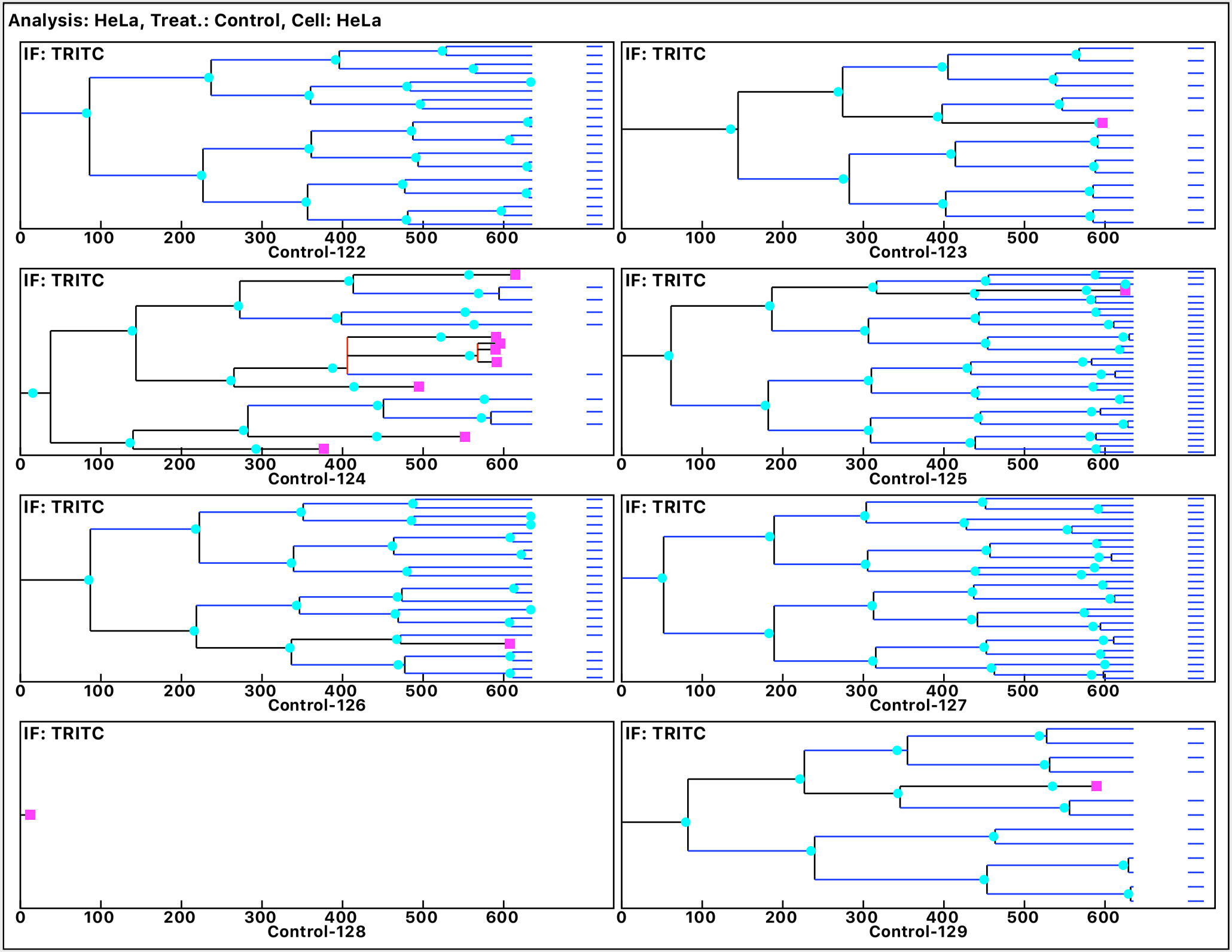
Analysis: HeLa, Treat.: Control, Cell: HeLa



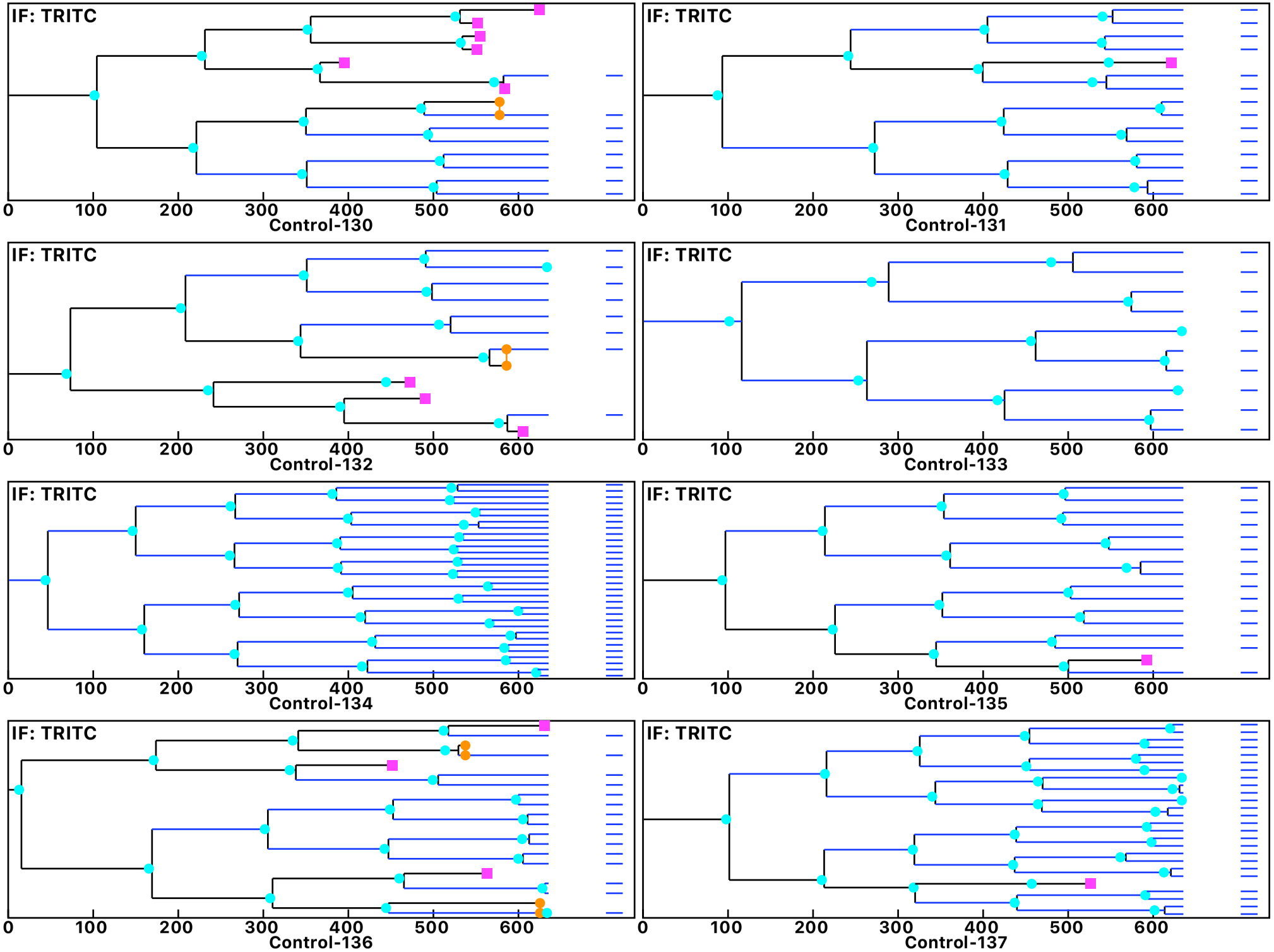
Analysis: HeLa, Treat.: Control, Cell: HeLa



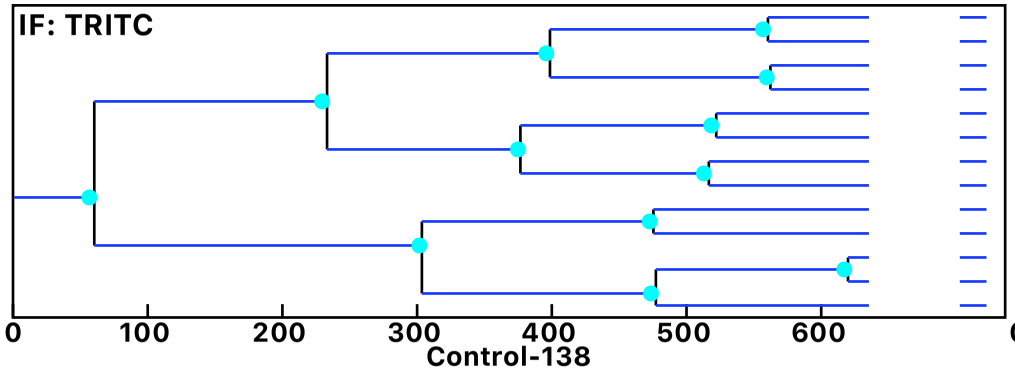
Analysis: HeLa, Treat.: Control, Cell: HeLa

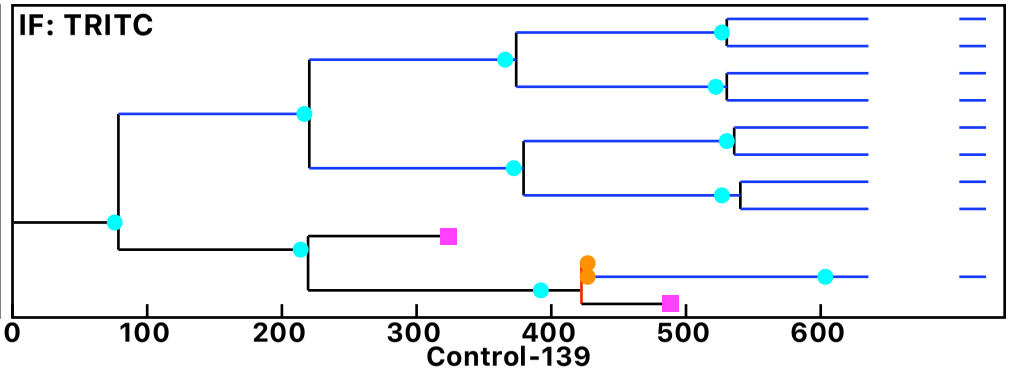


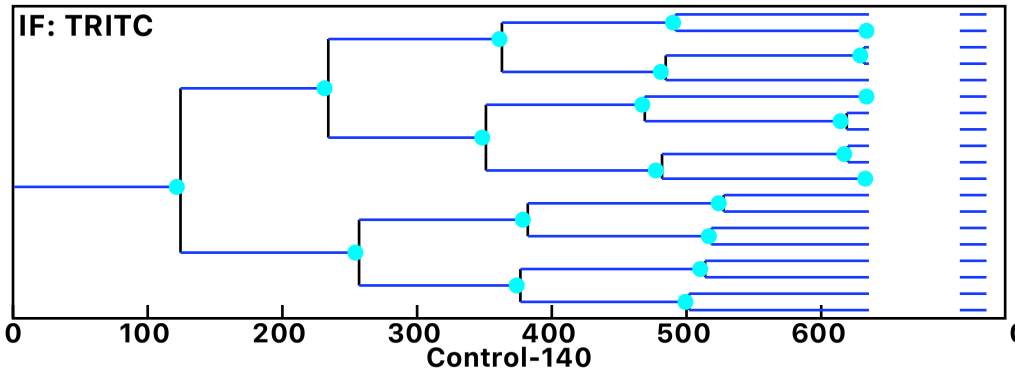
Analysis: HeLa, Treat.: Control, Cell: HeLa

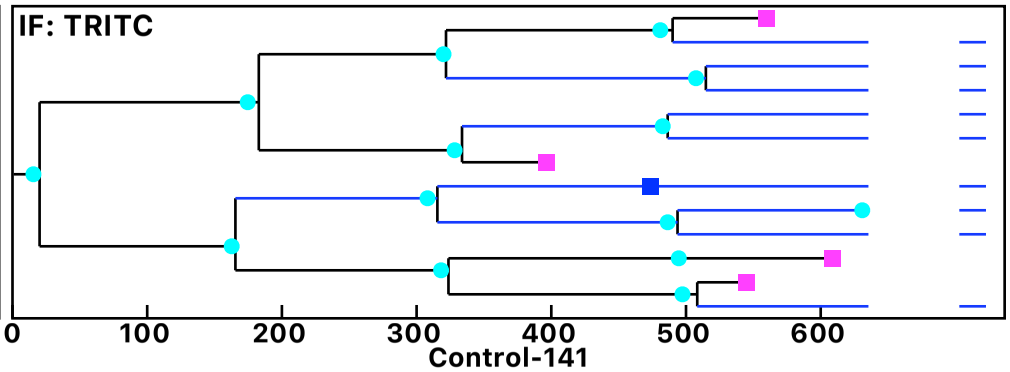


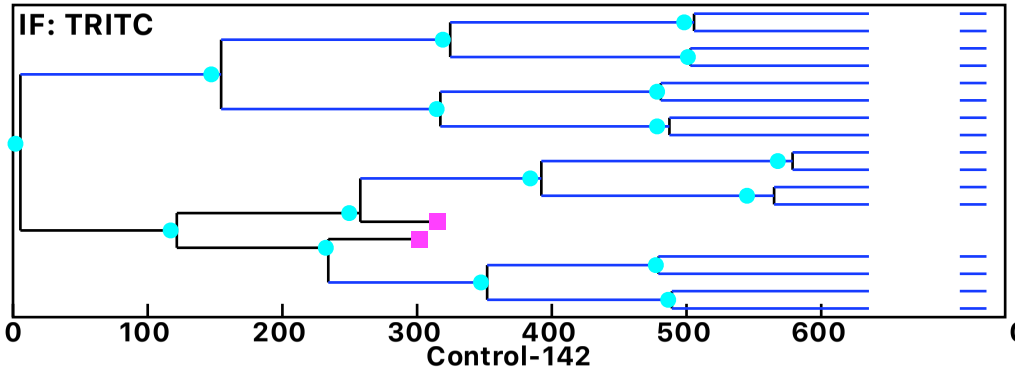
Analysis: HeLa, Treat.: Control, Cell: HeLa

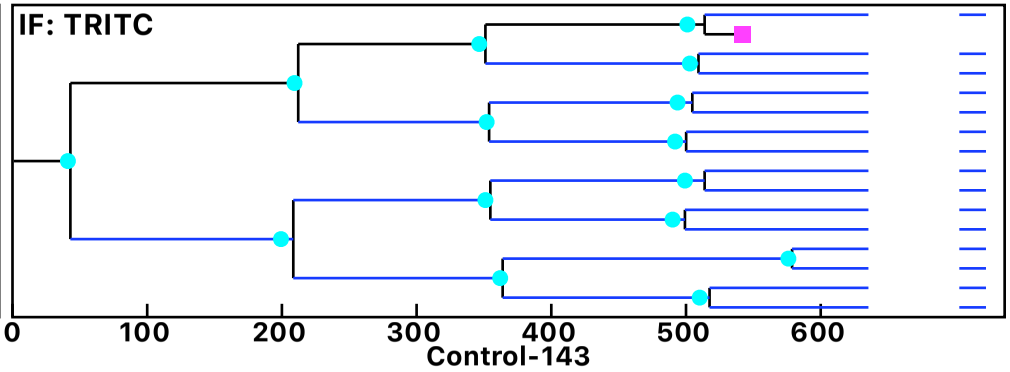


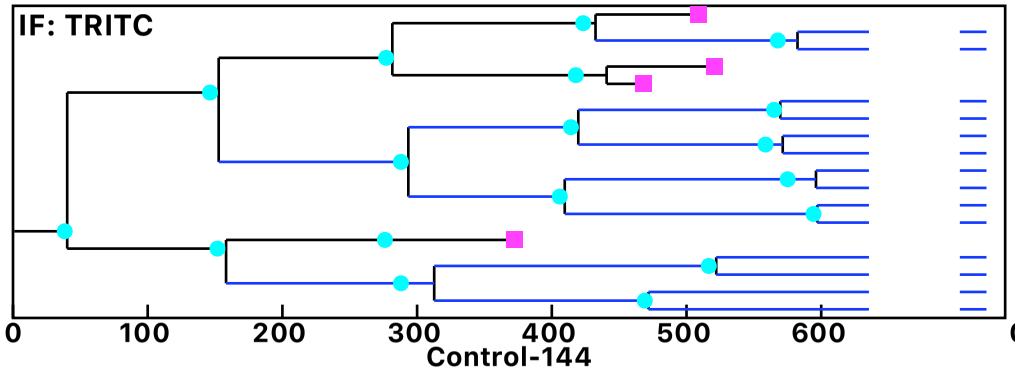


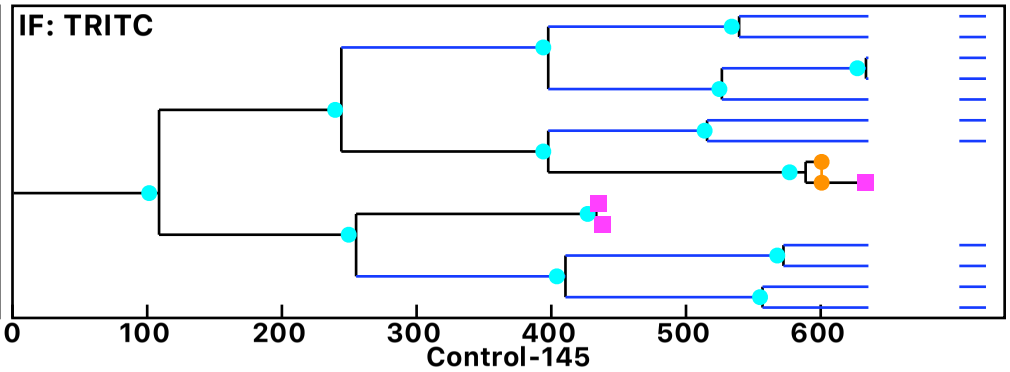




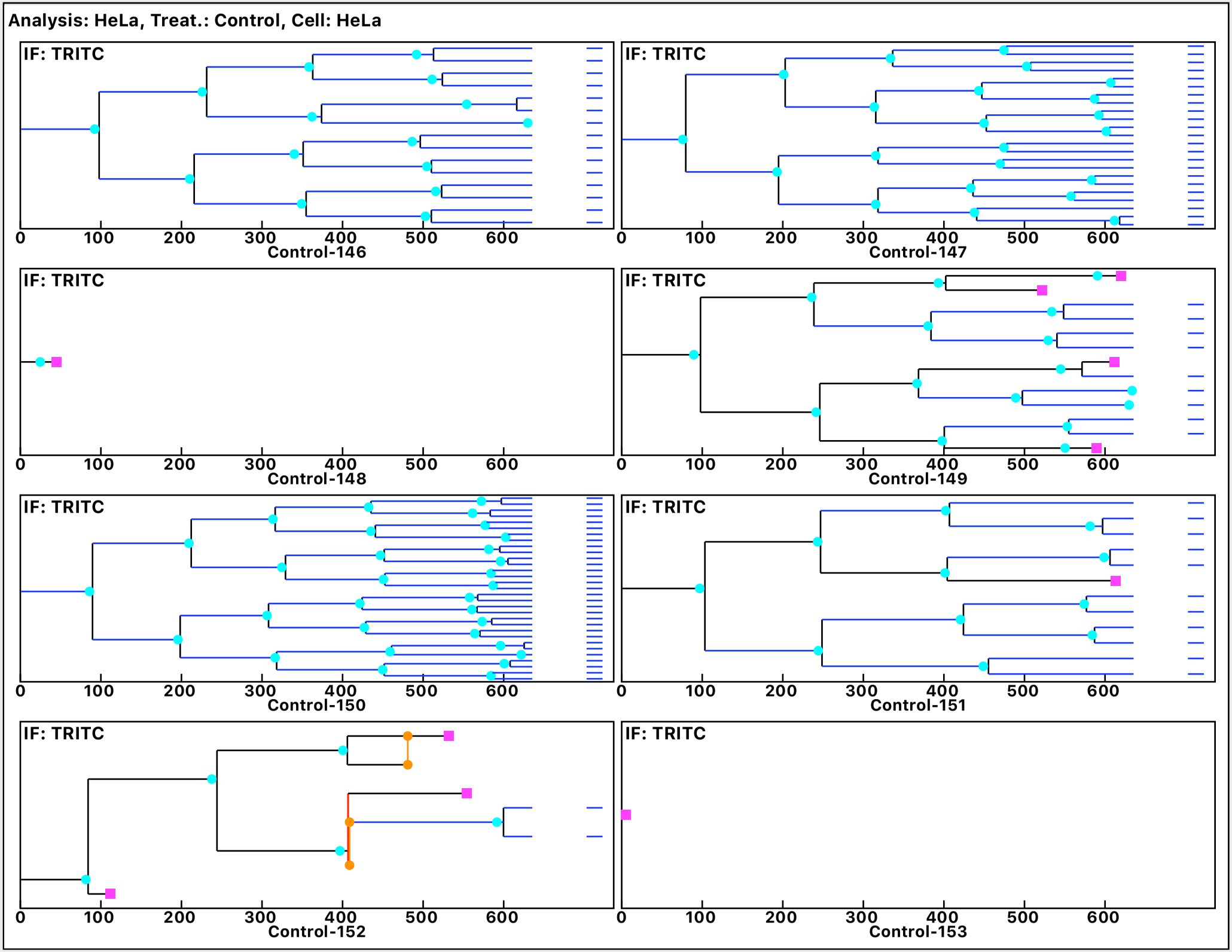




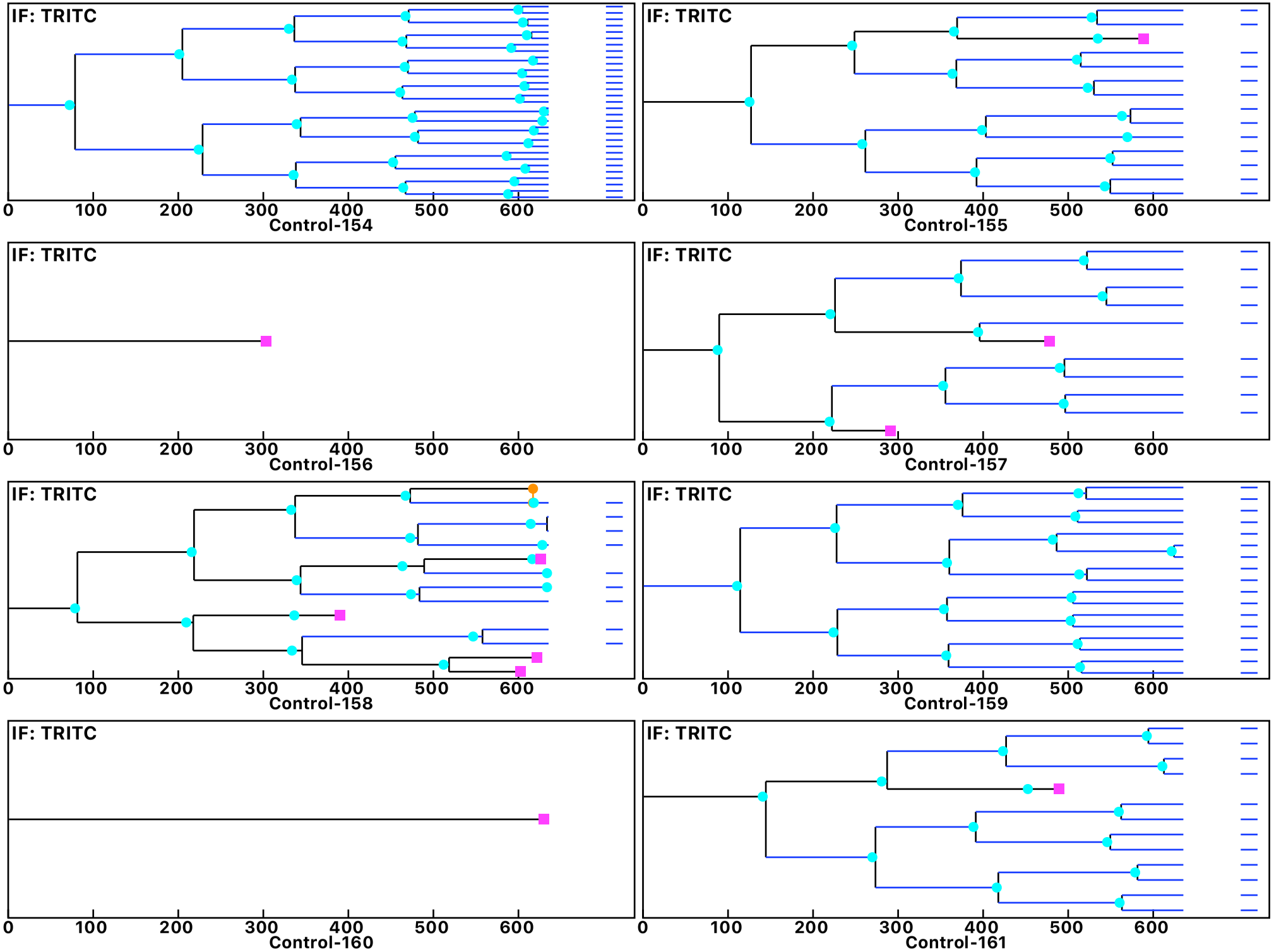




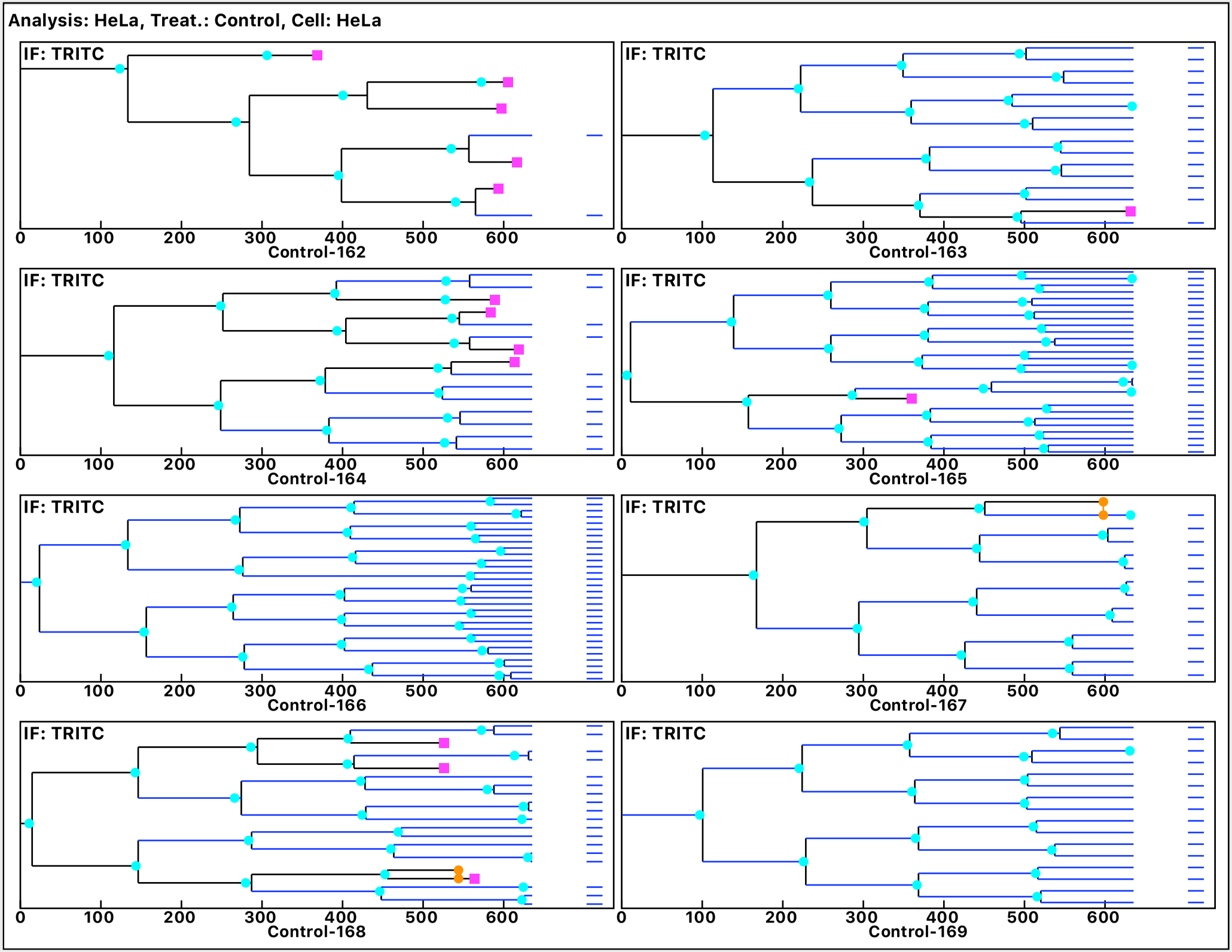
Analysis: HeLa, Treat.: Control, Cell: HeLa



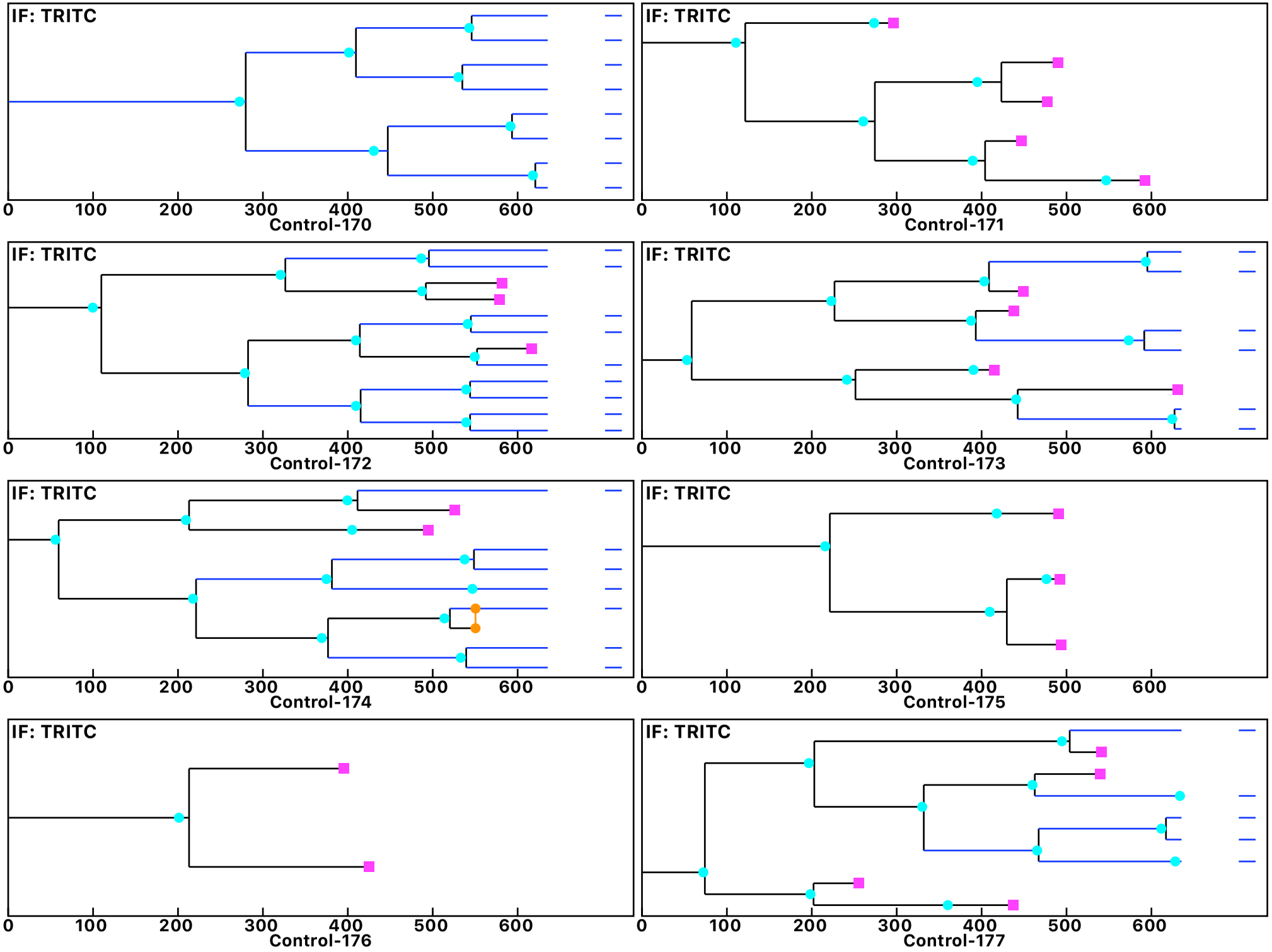
Analysis: HeLa, Treat.: Control, Cell: HeLa



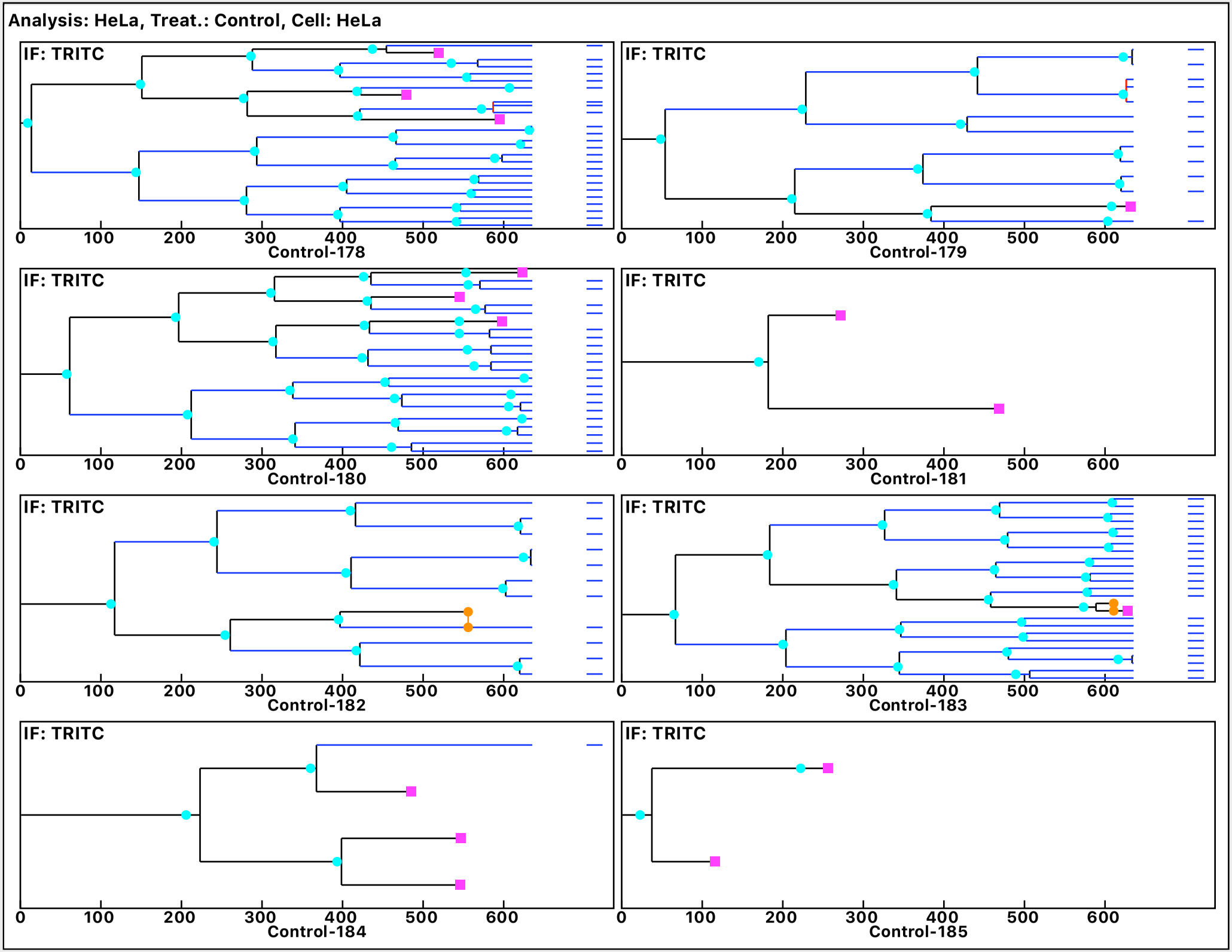
Analysis: HeLa, Treat.: Control, Cell: HeLa



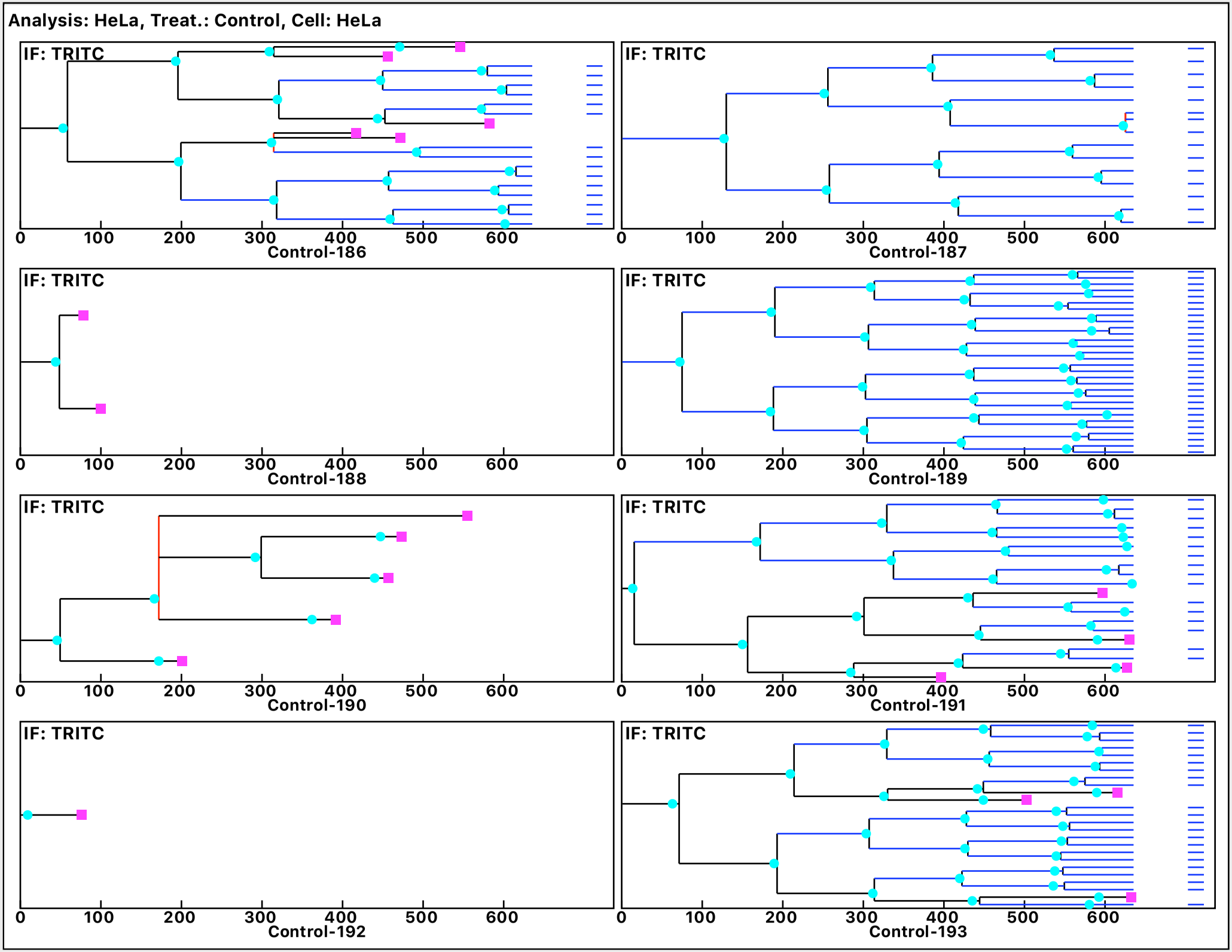
Analysis: HeLa, Treat.: Control, Cell: HeLa



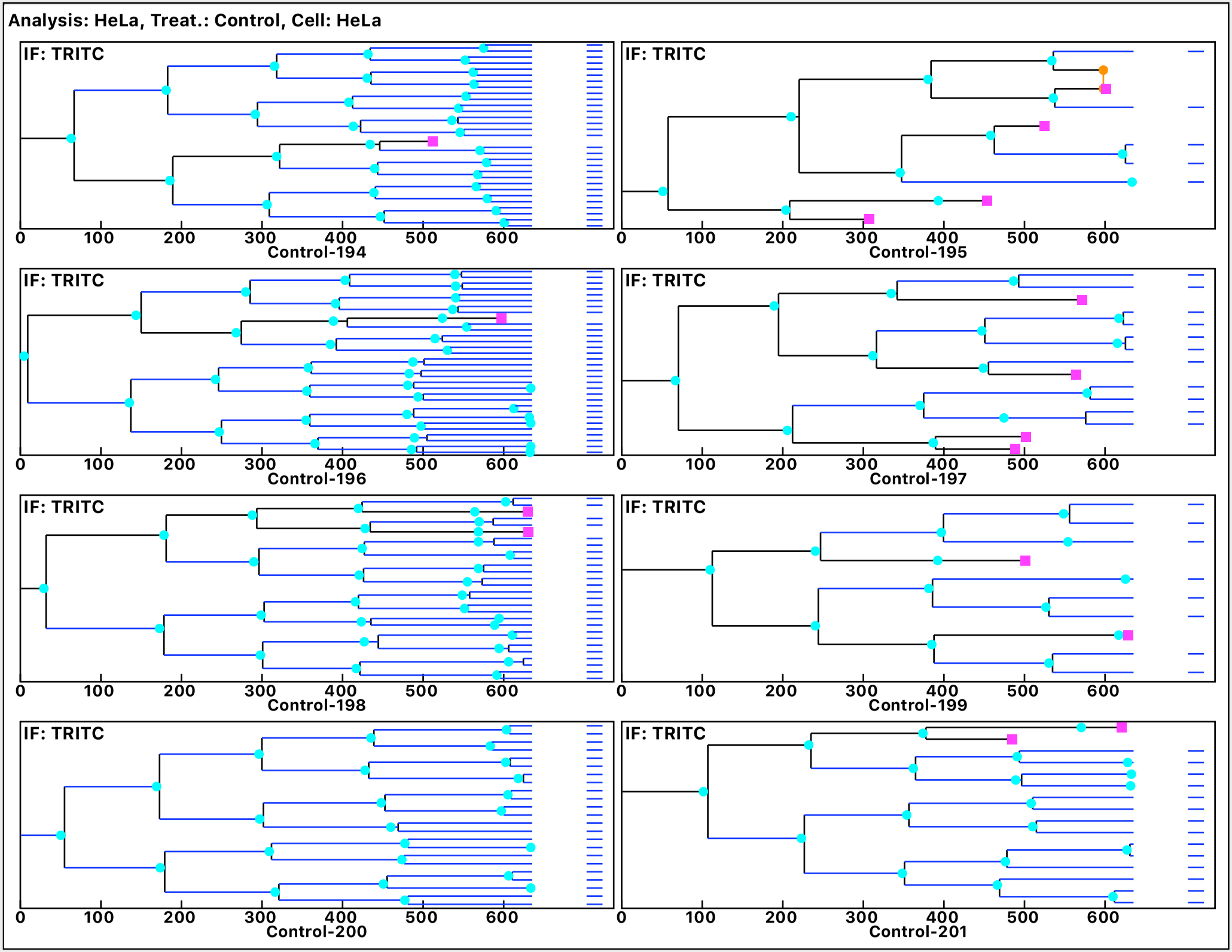
Analysis: HeLa, Treat.: Control, Cell: HeLa



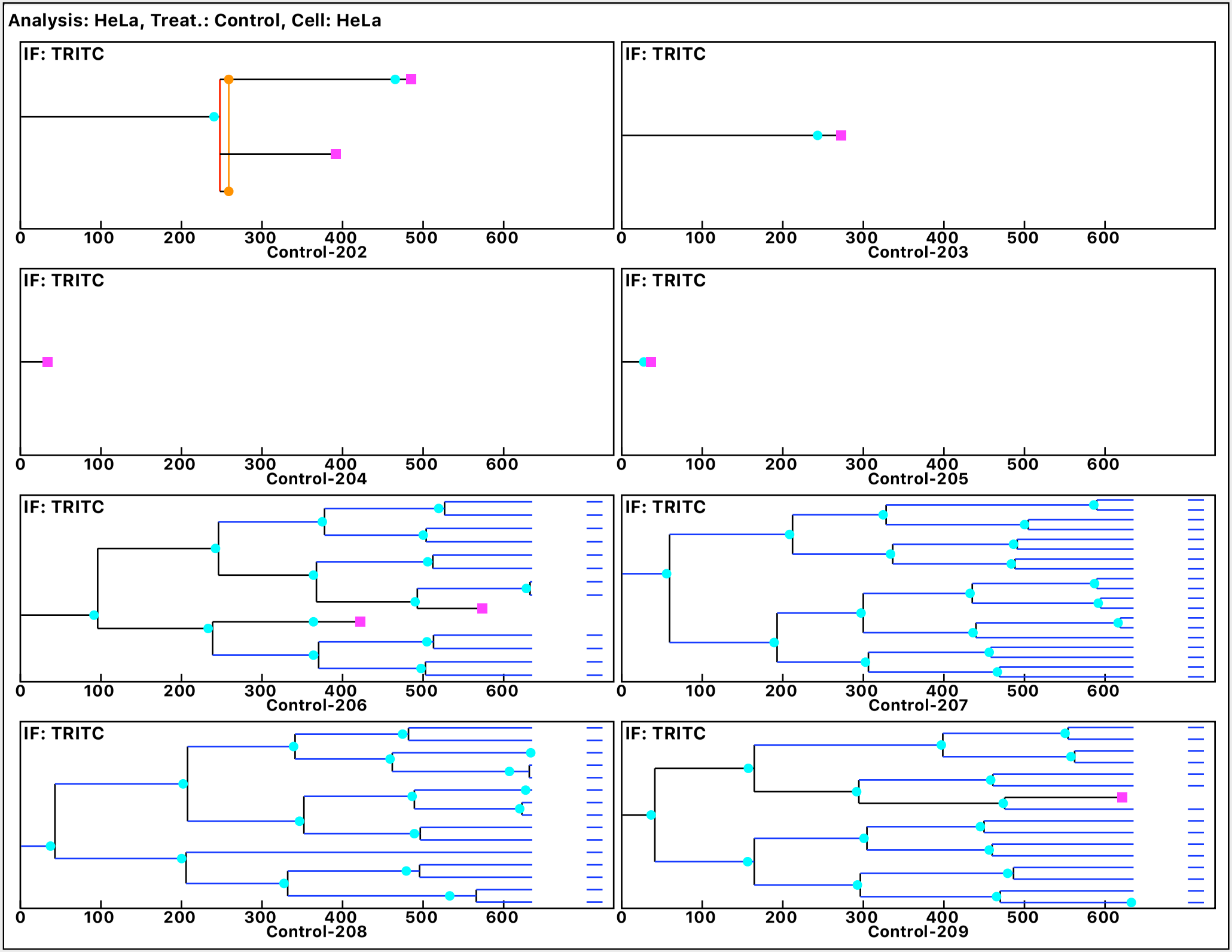
Analysis: HeLa, Treat.: Control, Cell: HeLa



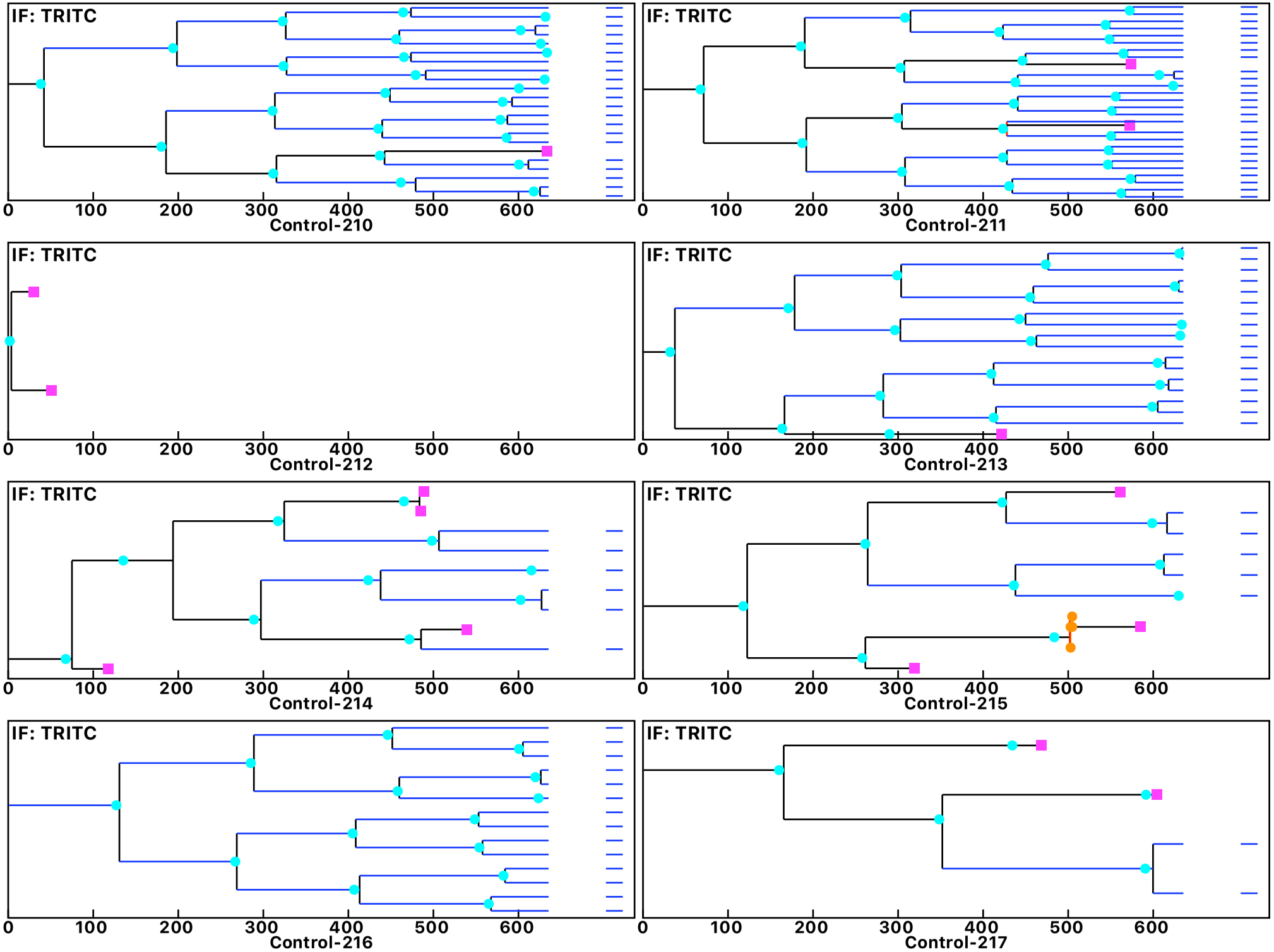
Analysis: HeLa, Treat.: Control, Cell: HeLa



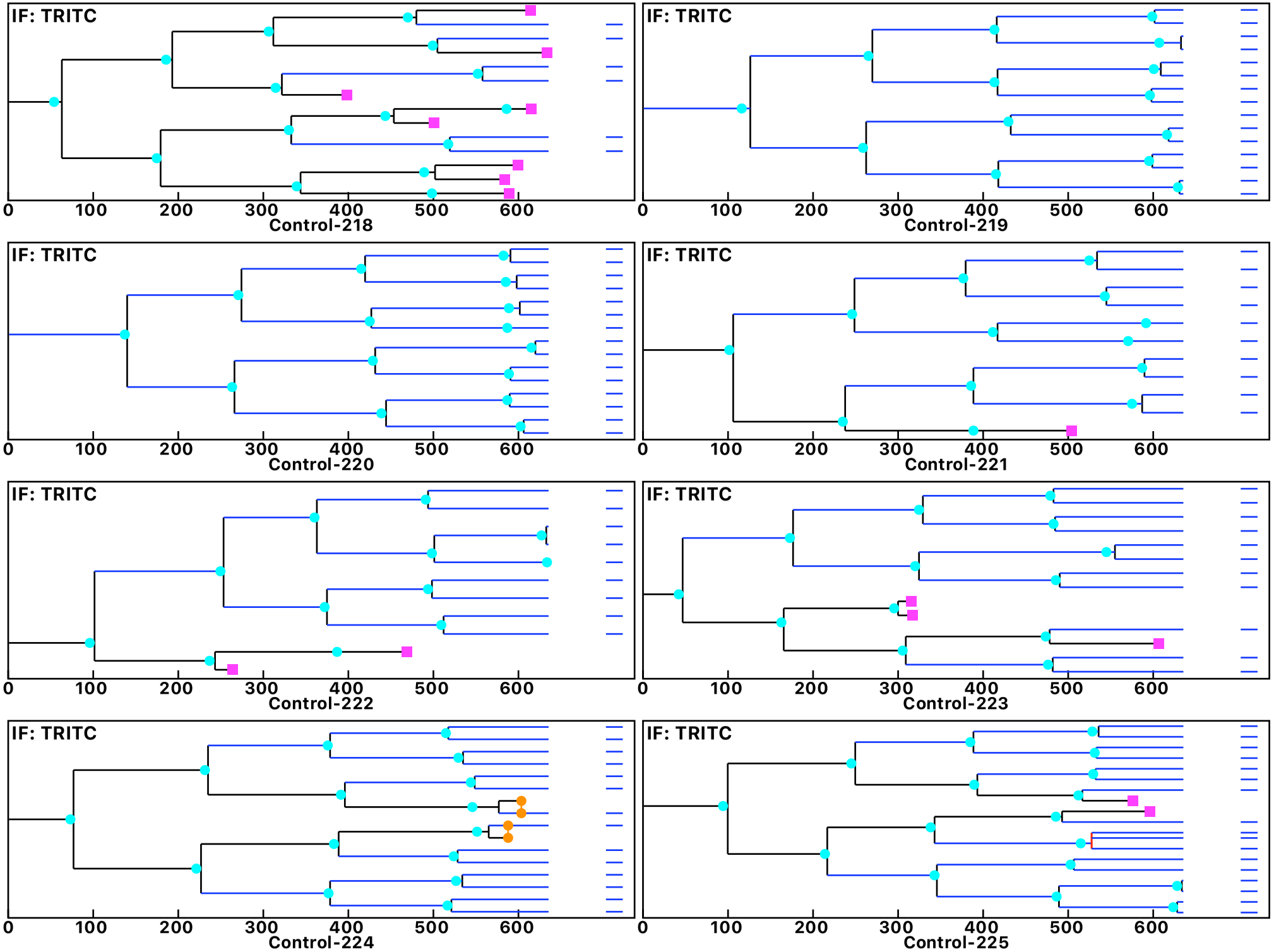
Analysis: HeLa, Treat.: Control, Cell: HeLa



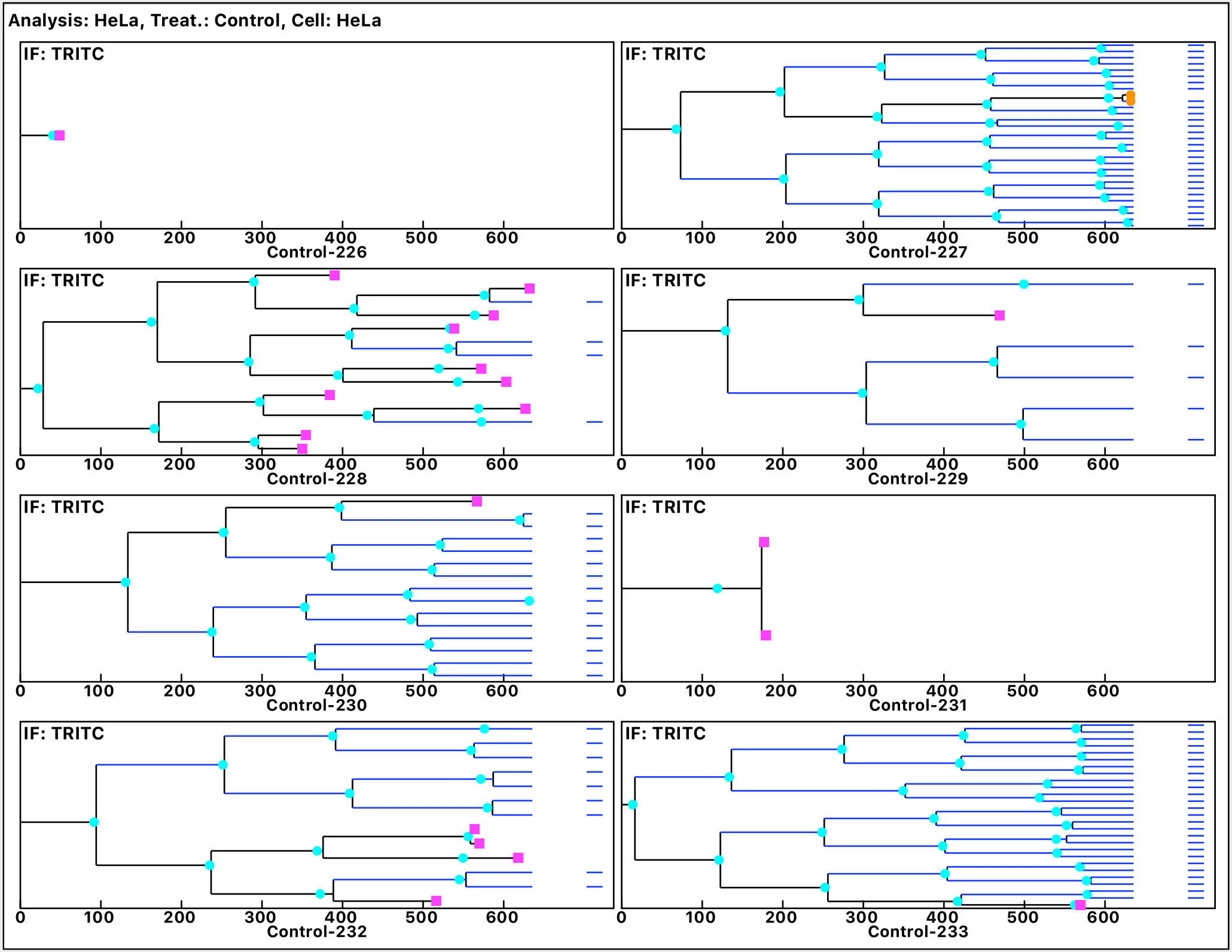
Analysis: HeLa, Treat.: Control, Cell: HeLa



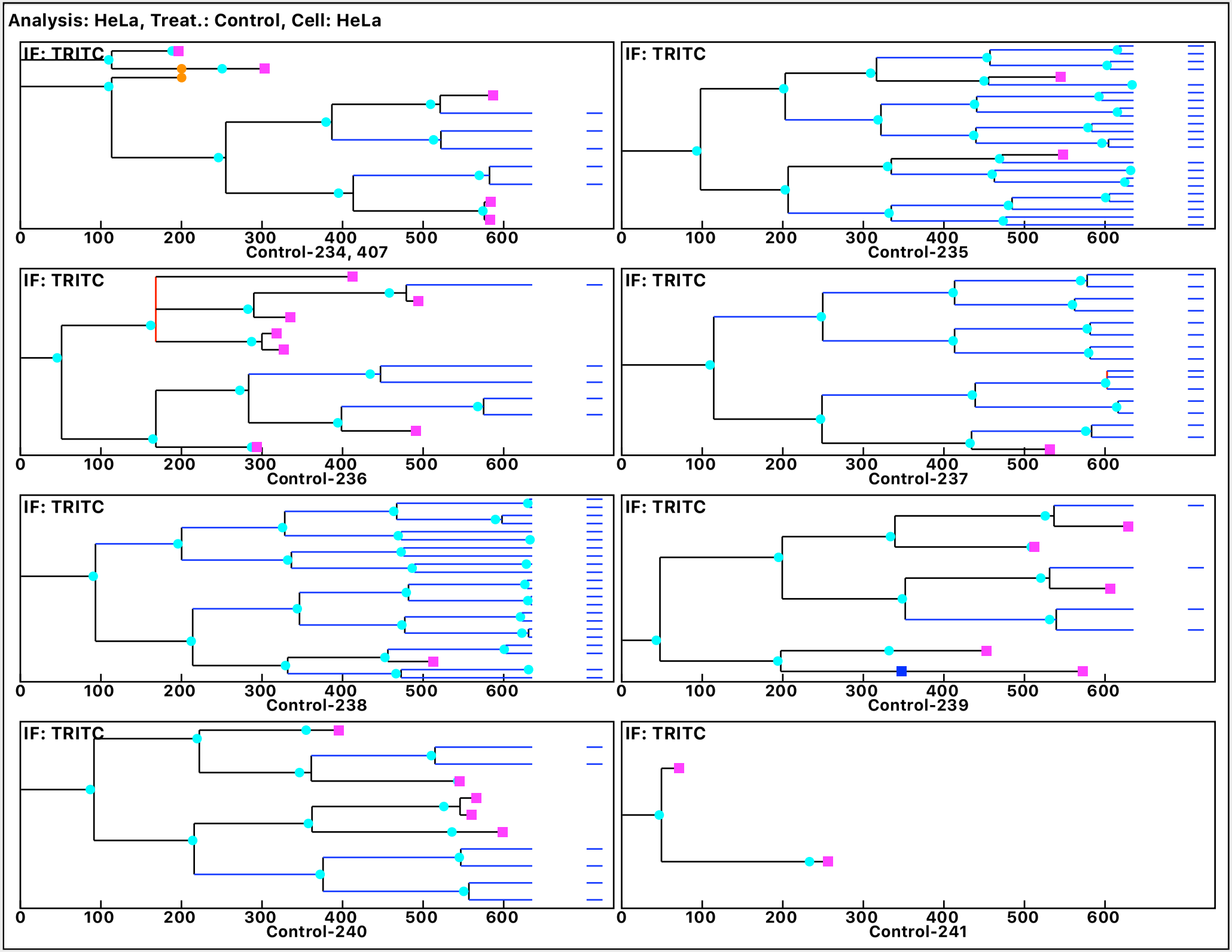
Analysis: HeLa, Treat.: Control, Cell: HeLa



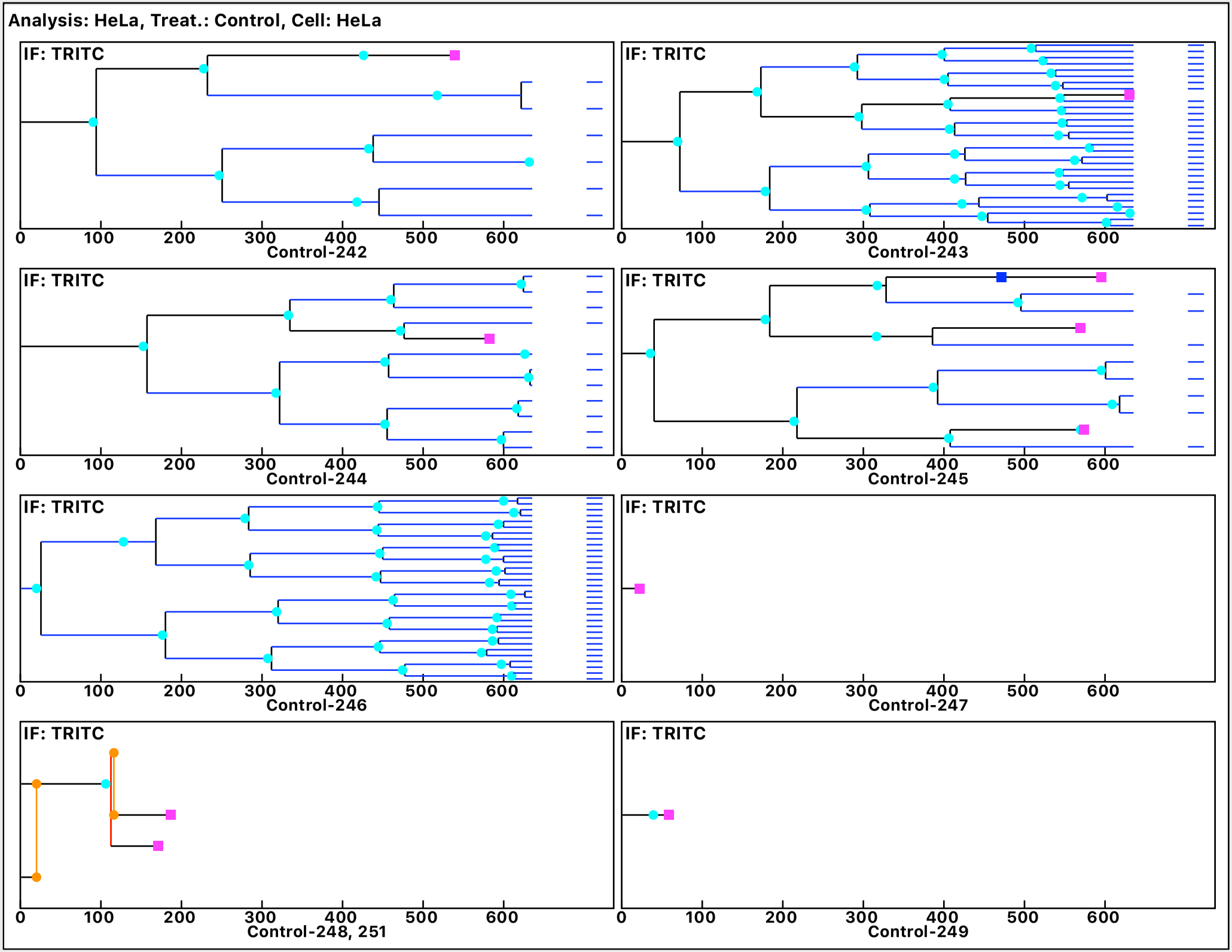
Analysis: HeLa, Treat.: Control, Cell: HeLa



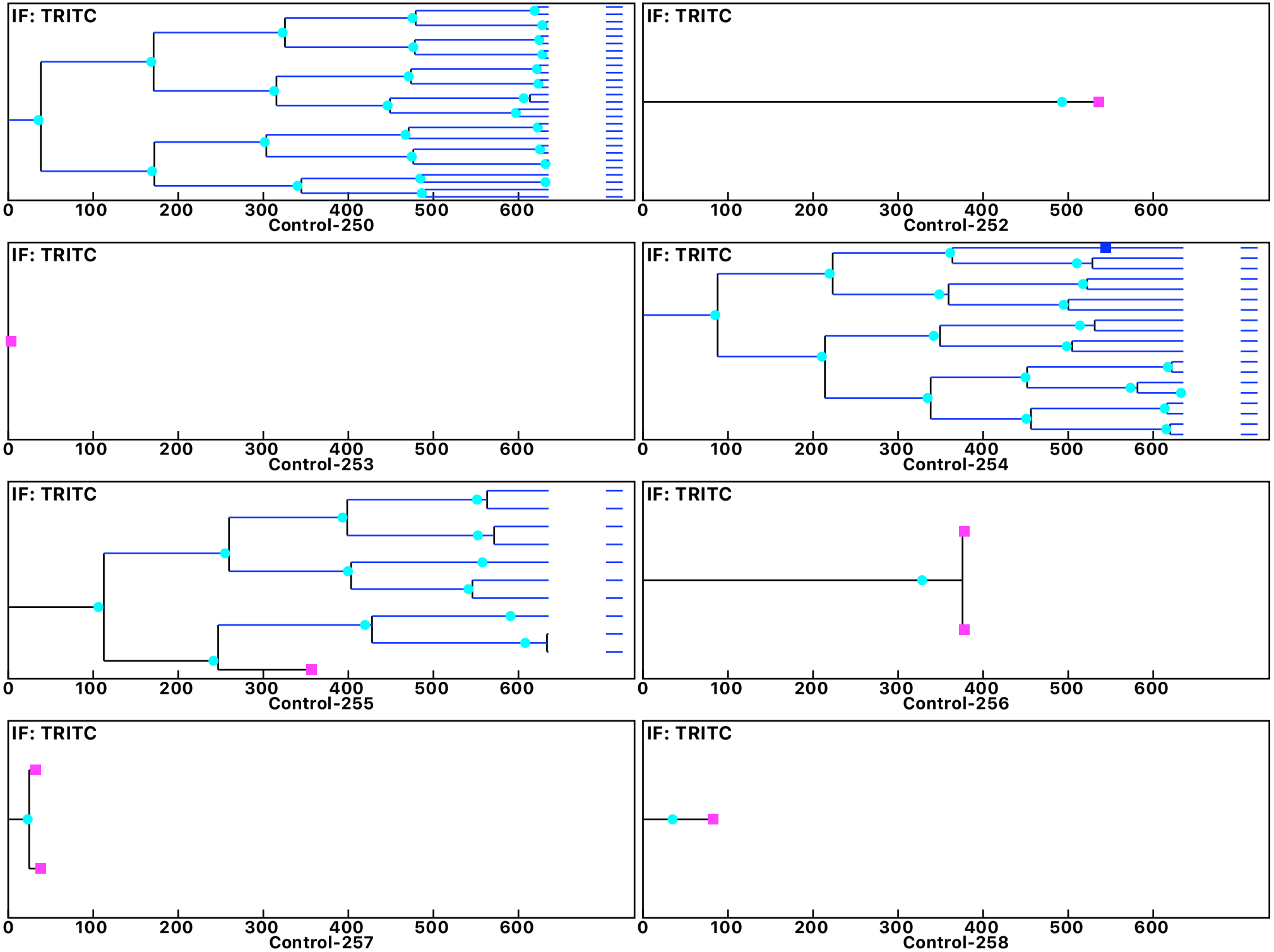
Analysis: HeLa, Treat.: Control, Cell: HeLa



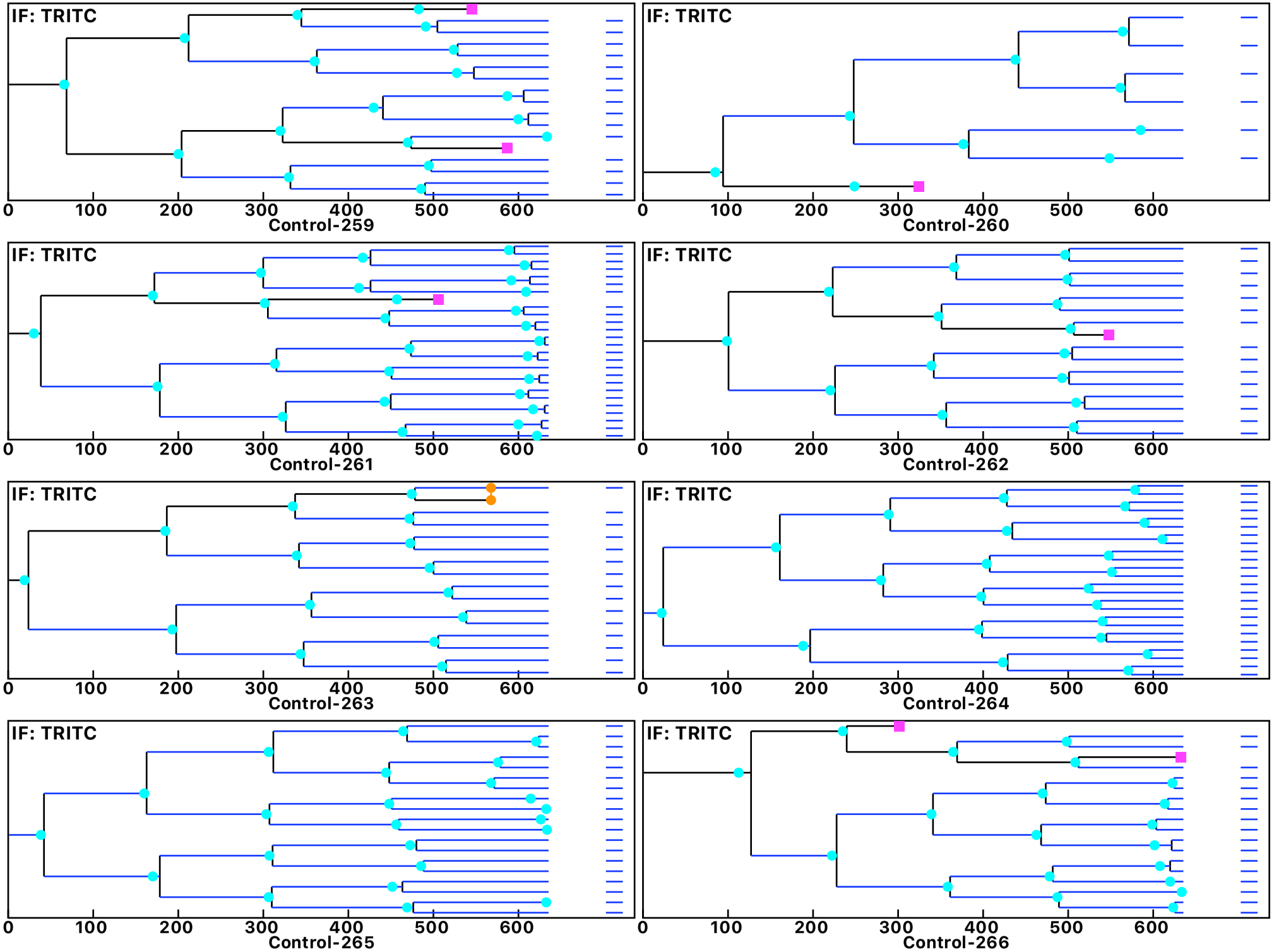
Analysis: HeLa, Treat.: Control, Cell: HeLa



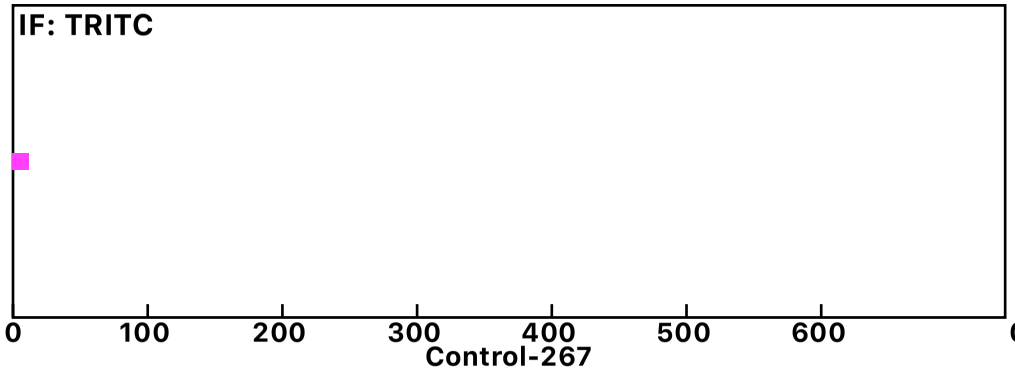
Analysis: HeLa, Treat.: Control, Cell: HeLa

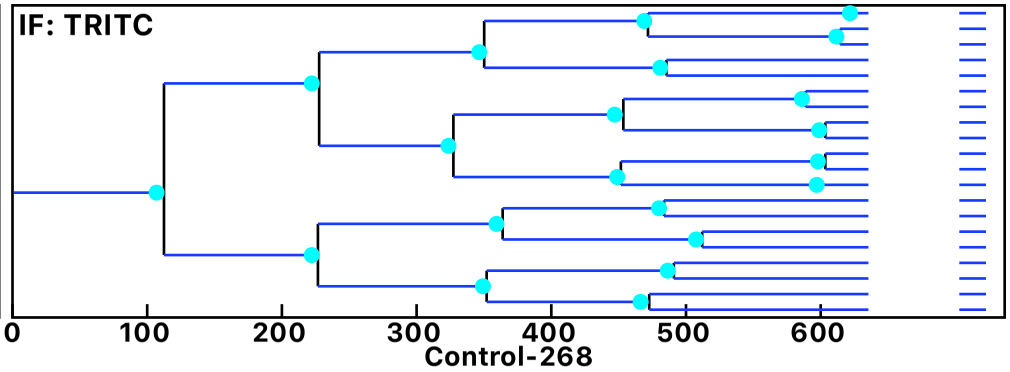


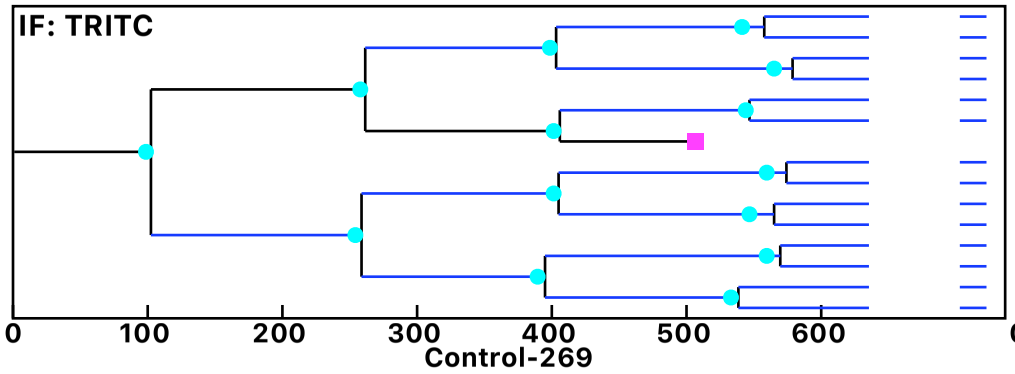
Analysis: HeLa, Treat.: Control, Cell: HeLa

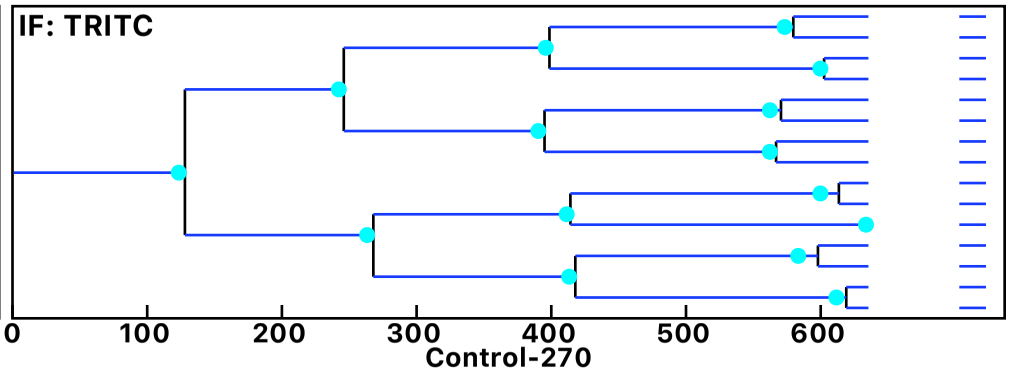


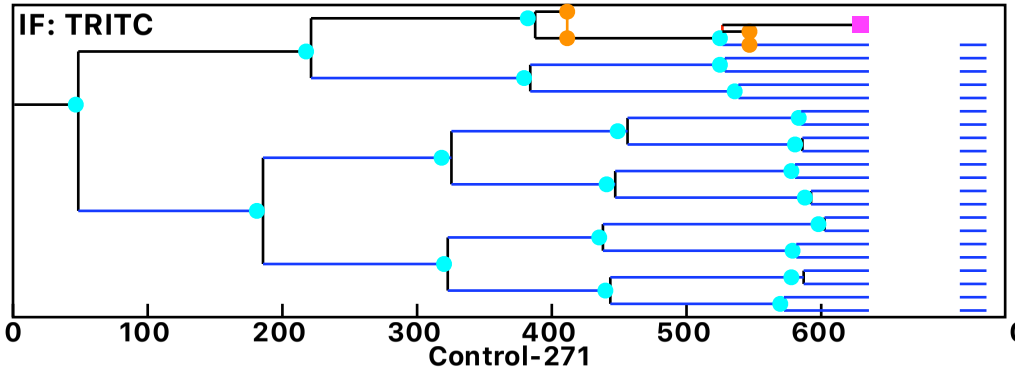
Analysis: HeLa, Treat.: Control, Cell: HeLa

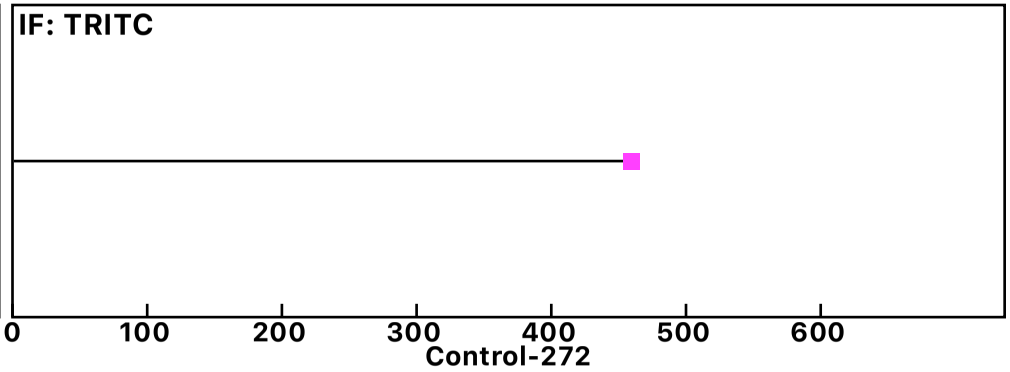


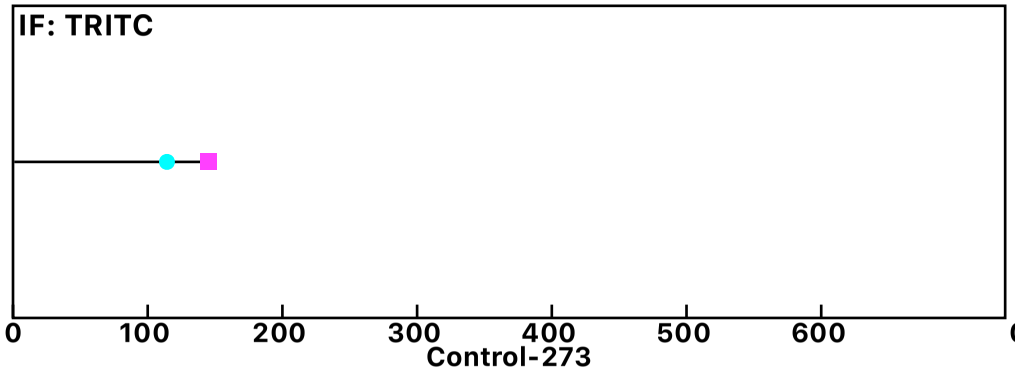


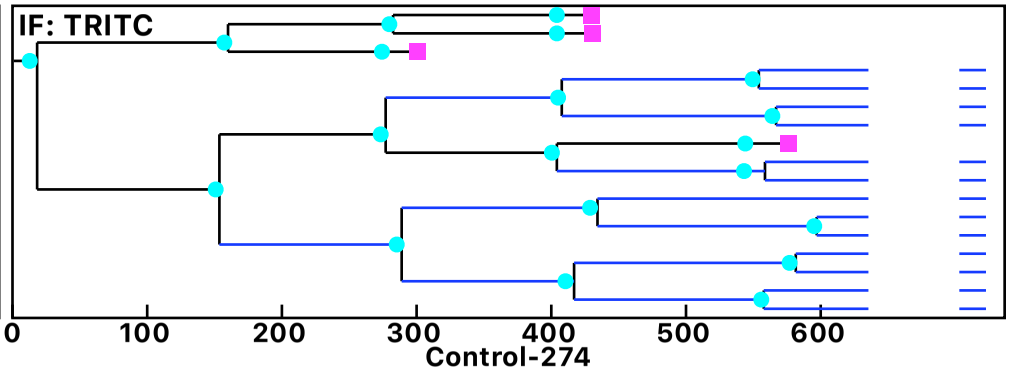




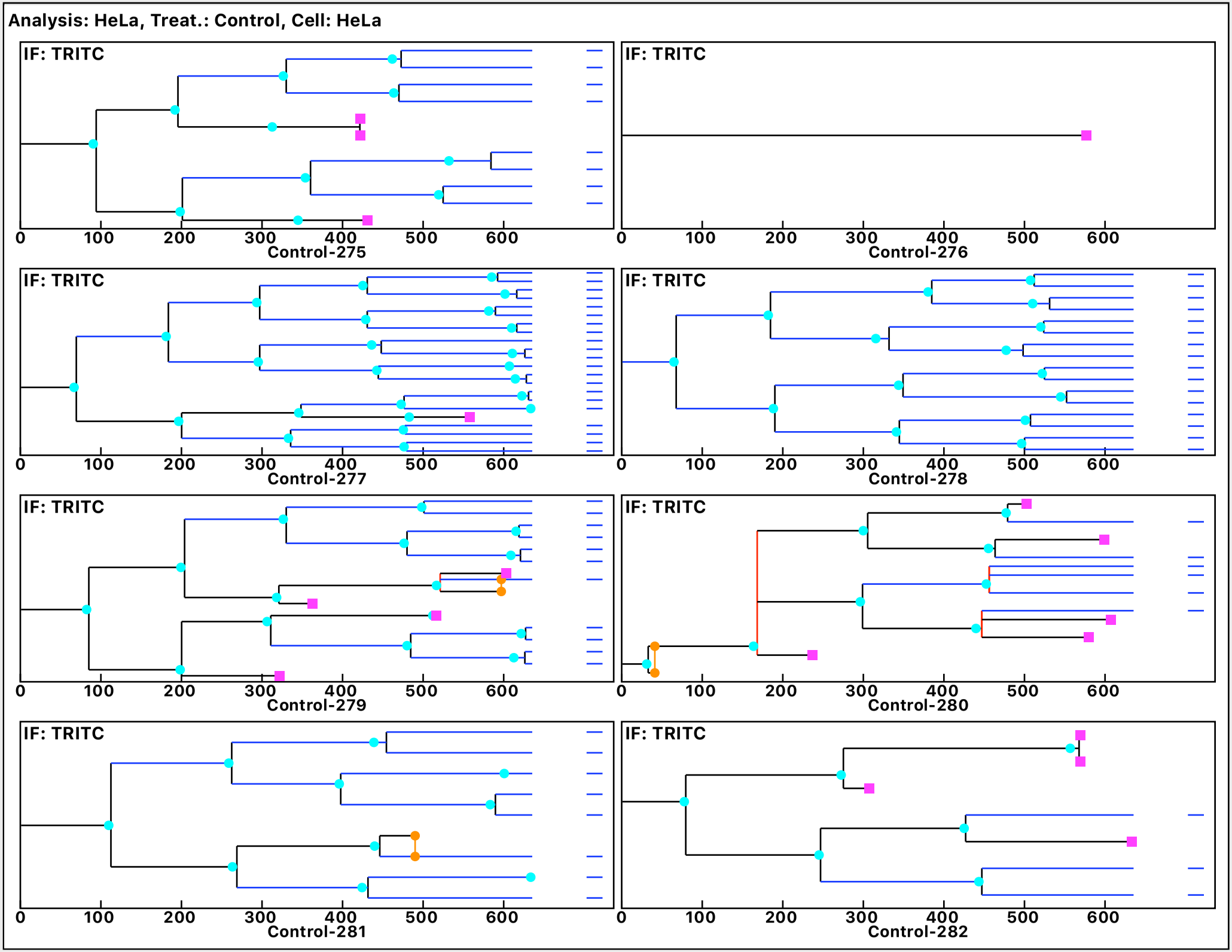




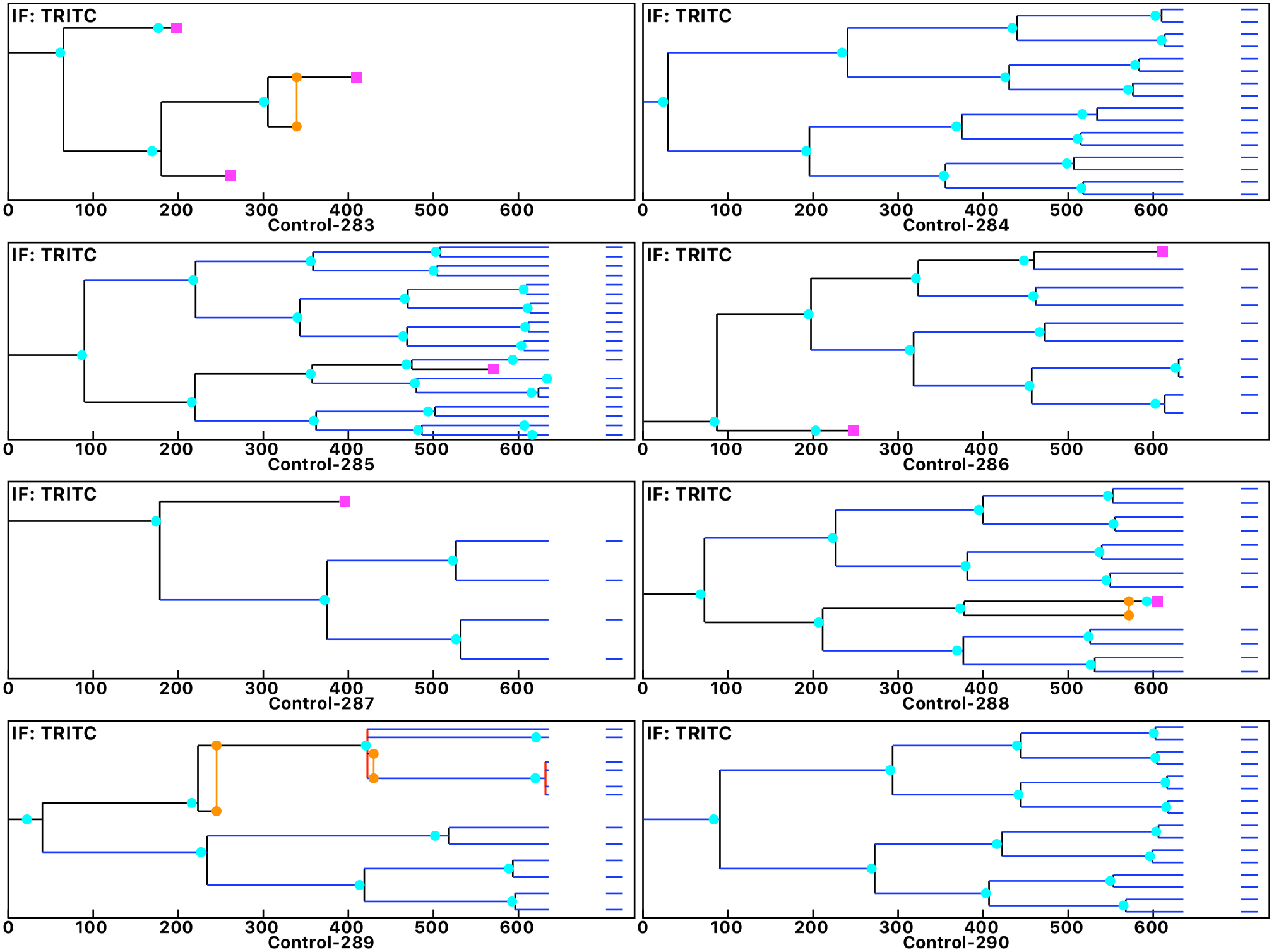




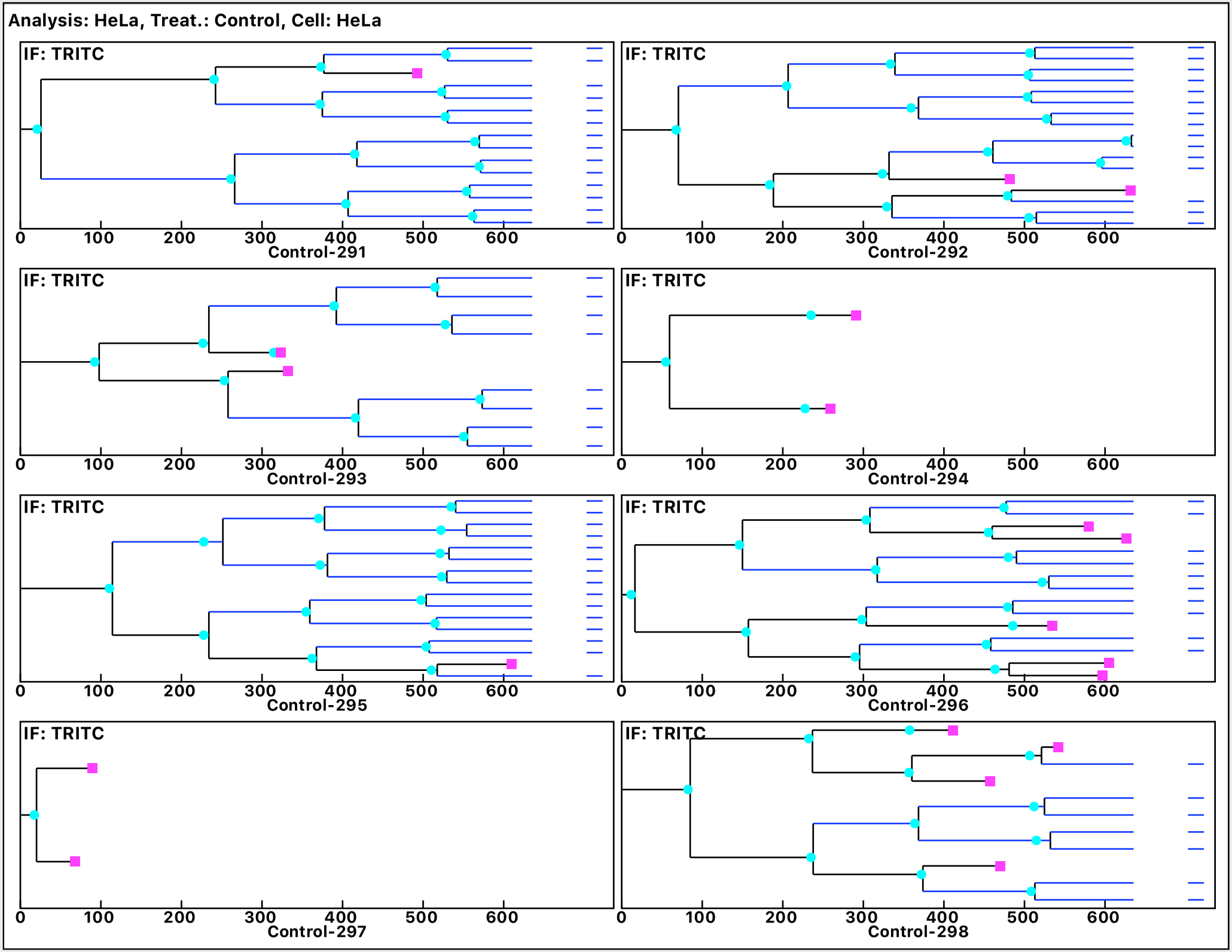
Analysis: HeLa, Treat.: Control, Cell: HeLa



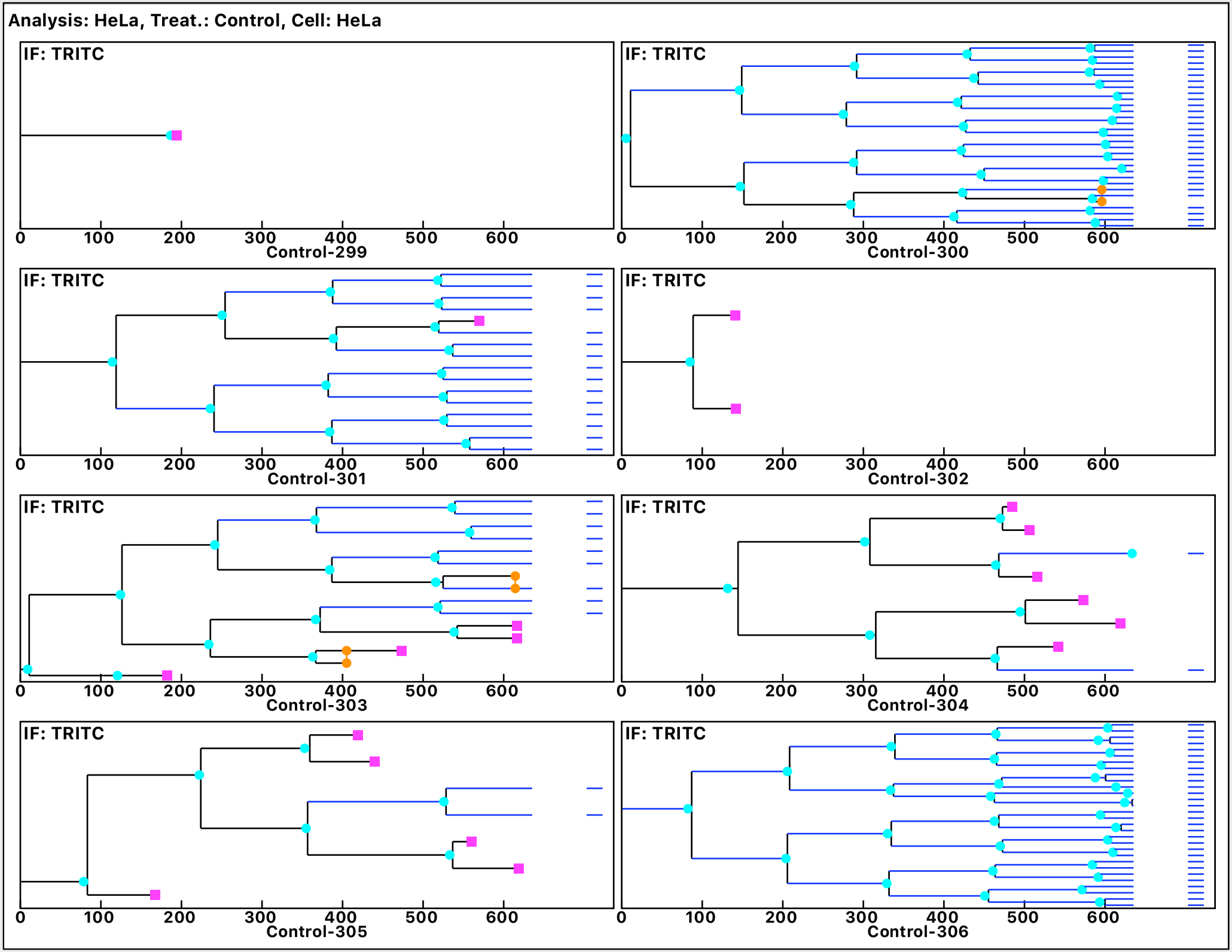
Analysis: HeLa, Treat.: Control, Cell: HeLa



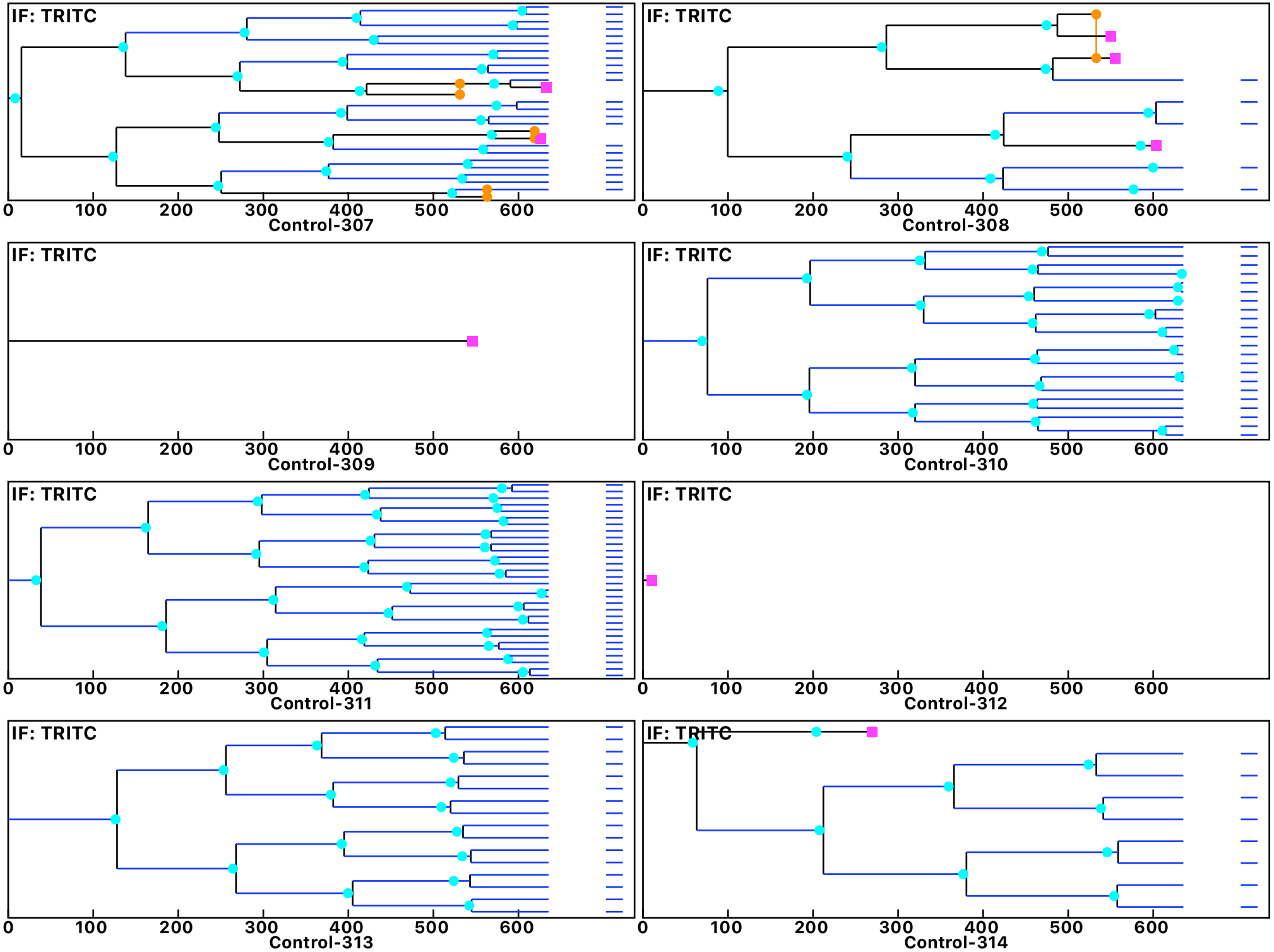
Analysis: HeLa, Treat.: Control, Cell: HeLa



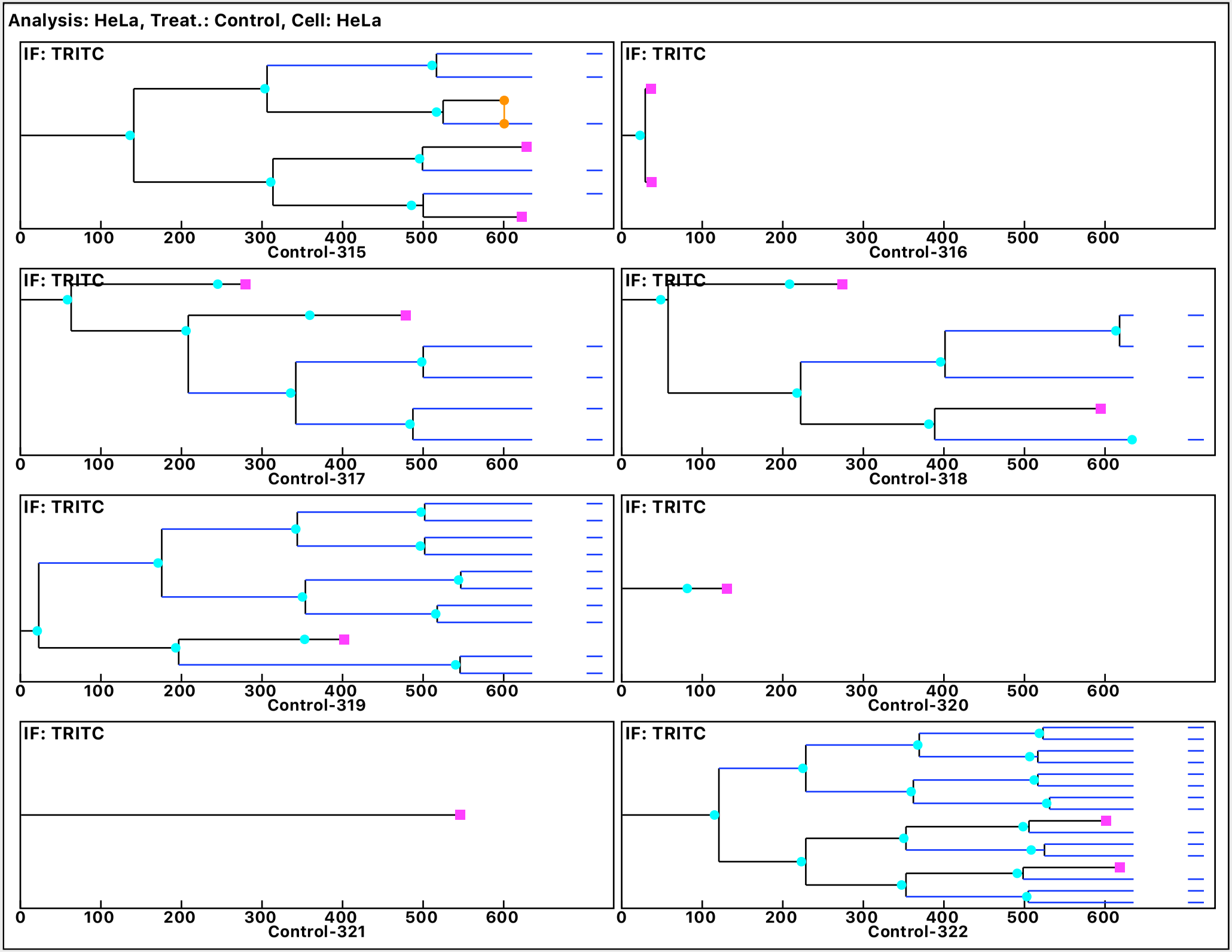
Analysis: HeLa, Treat.: Control, Cell: HeLa



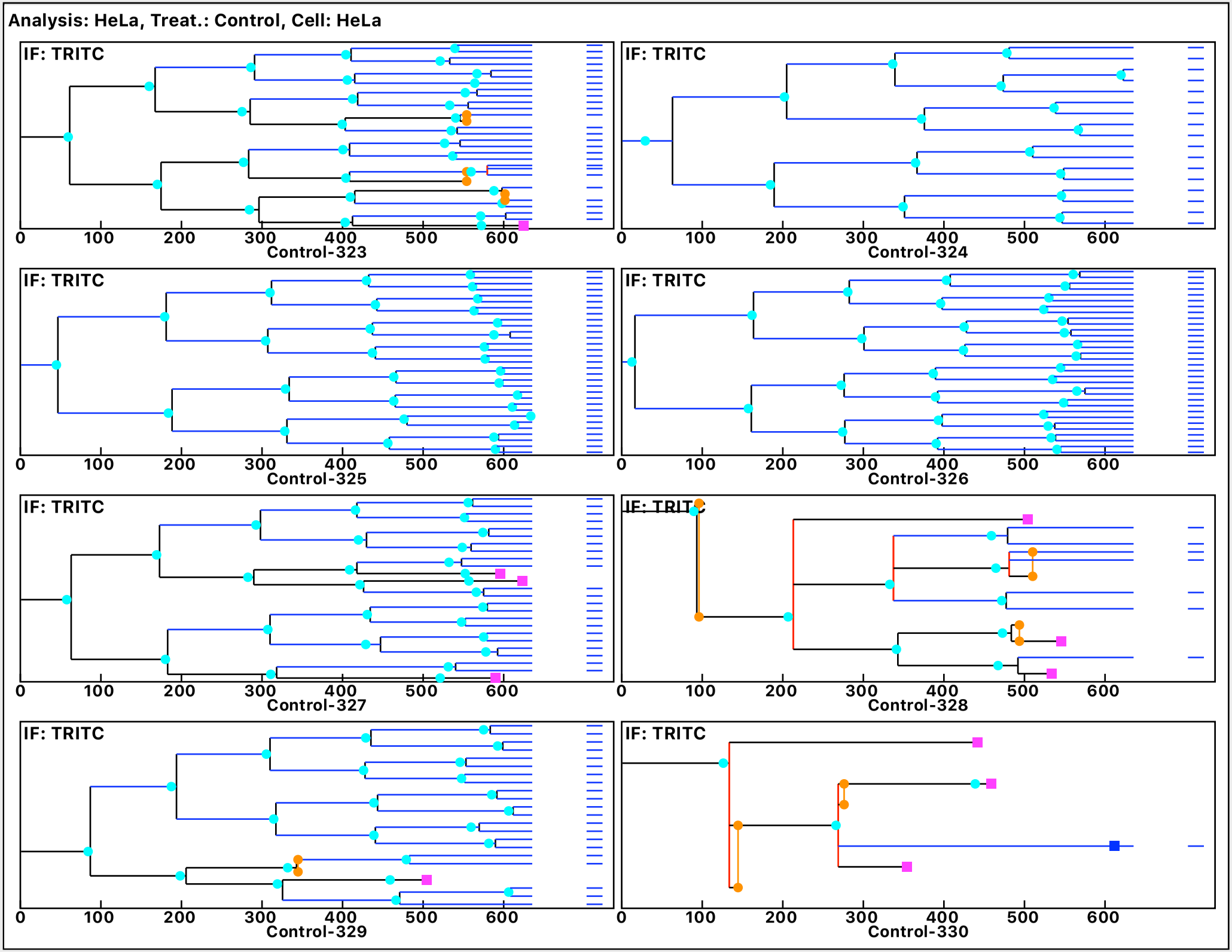
Analysis: HeLa, Treat.: Control, Cell: HeLa



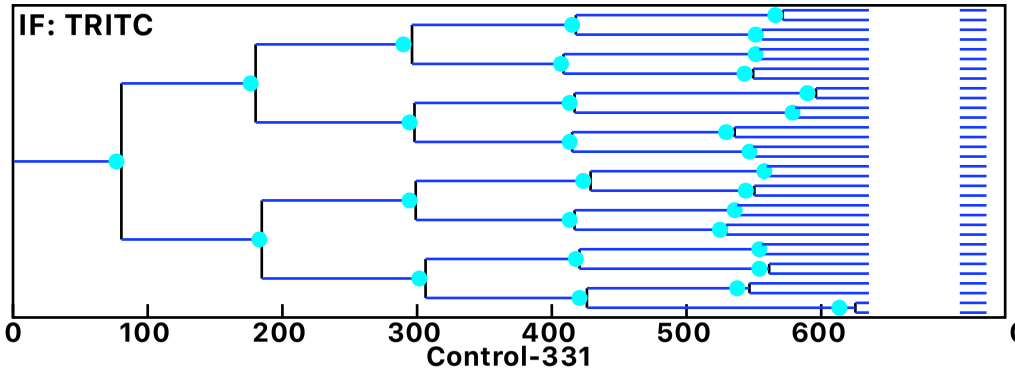
Analysis: HeLa, Treat.: Control, Cell: HeLa

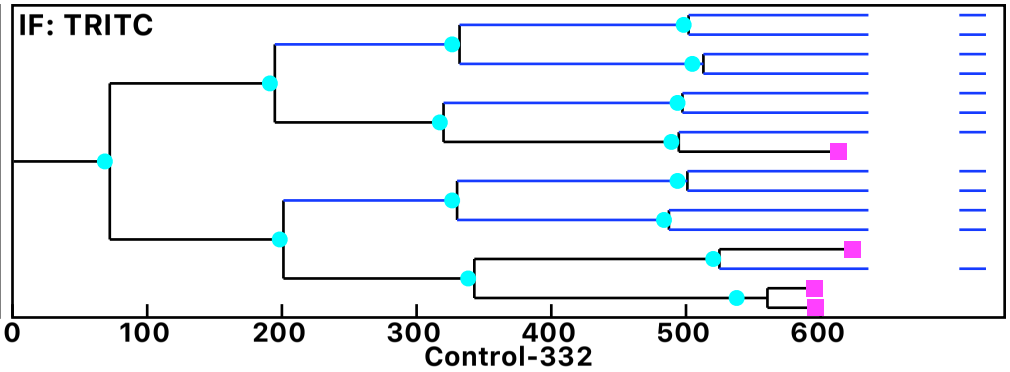


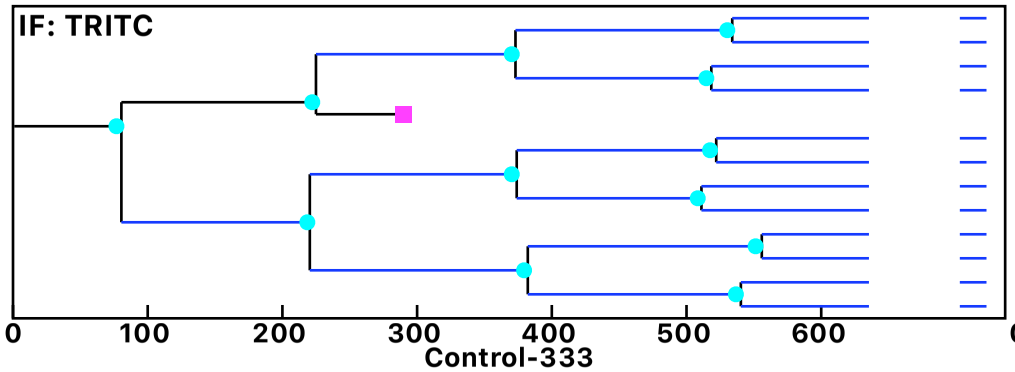
Analysis: HeLa, Treat.: Control, Cell: HeLa

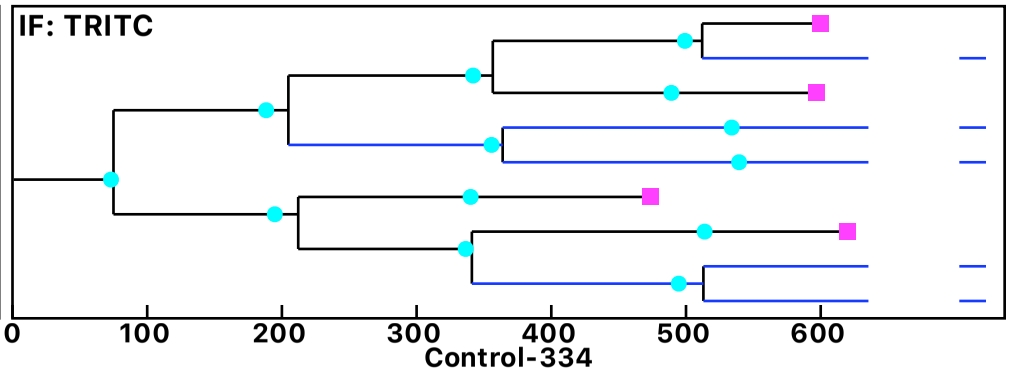


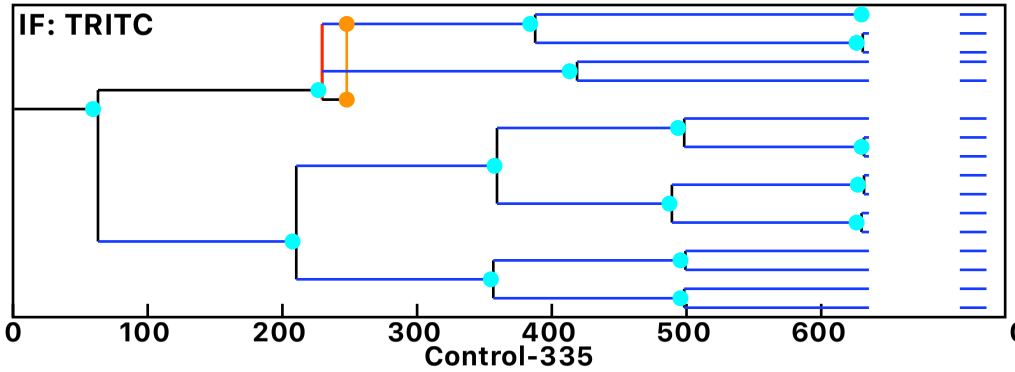
Analysis: HeLa, Treat.: Control, Cell: HeLa

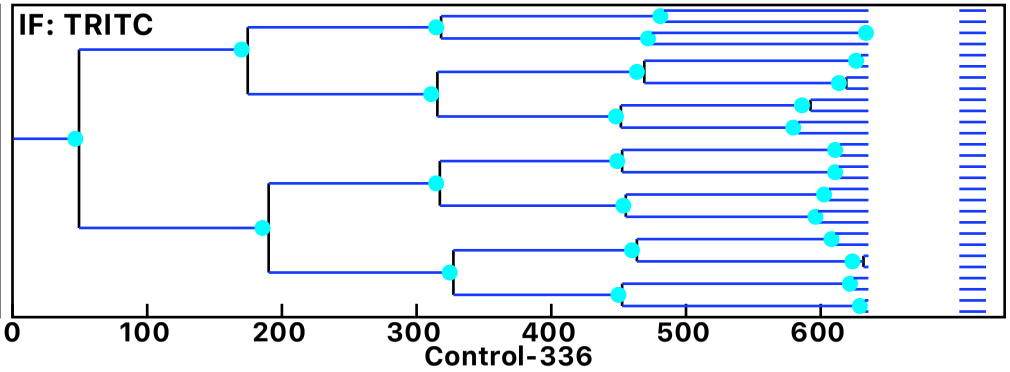


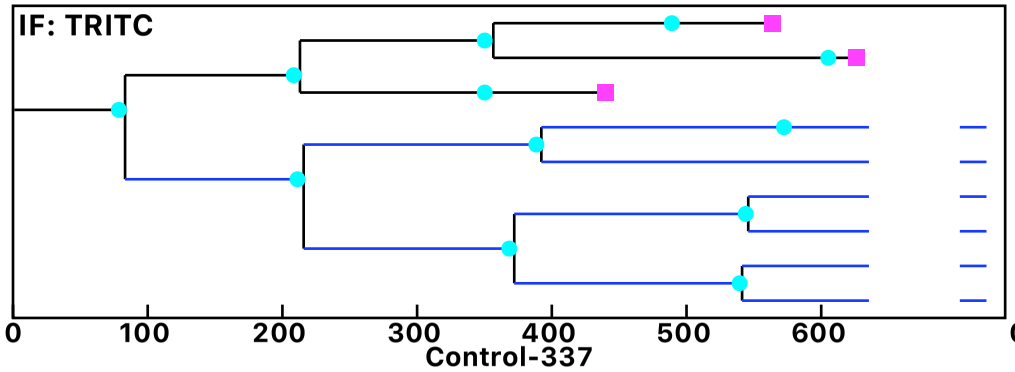


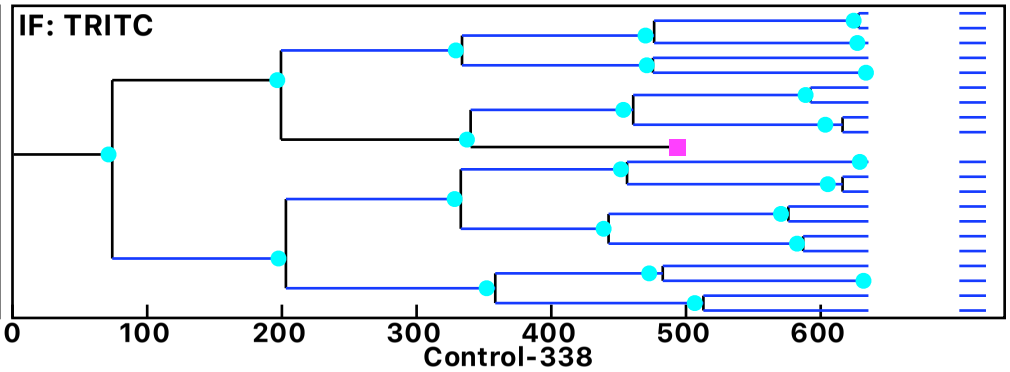




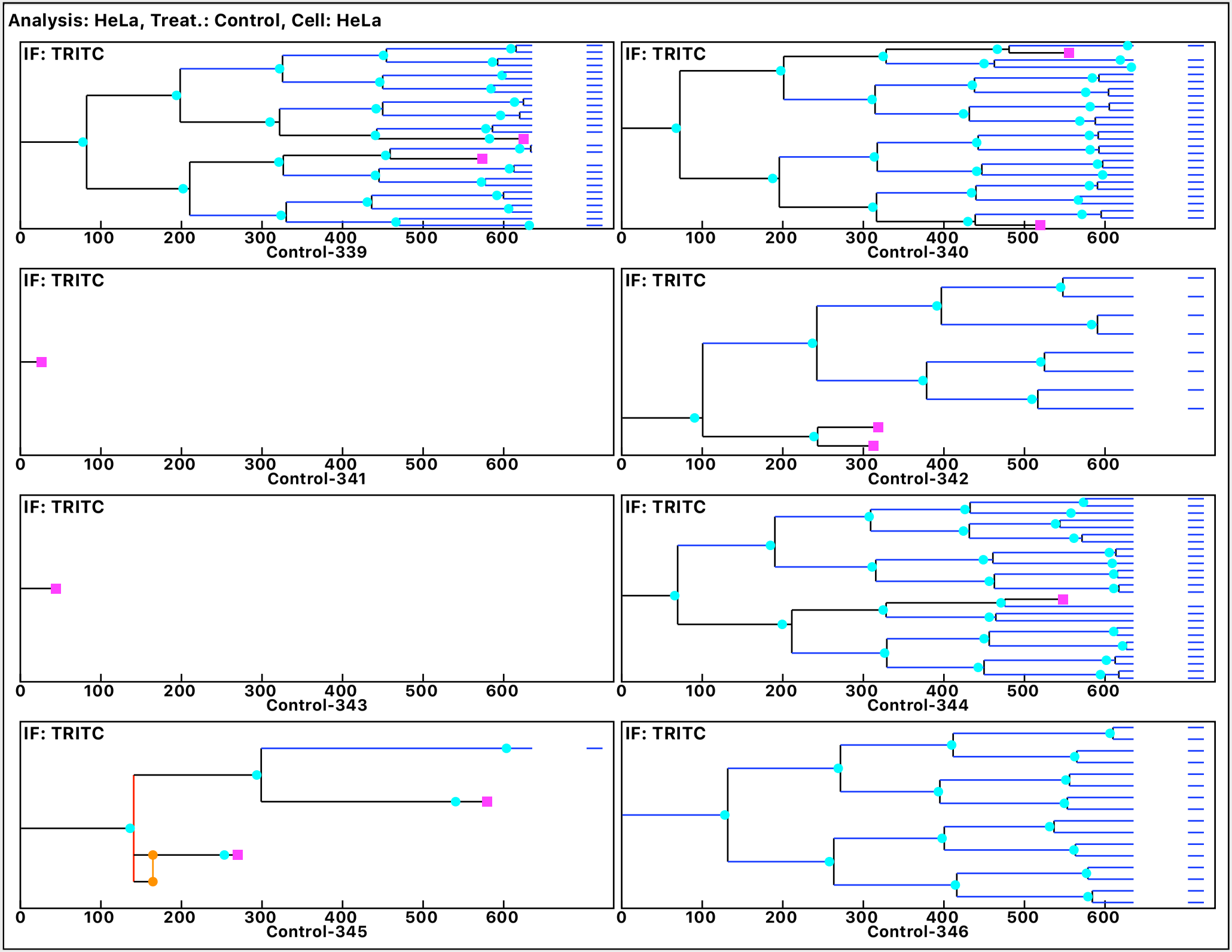




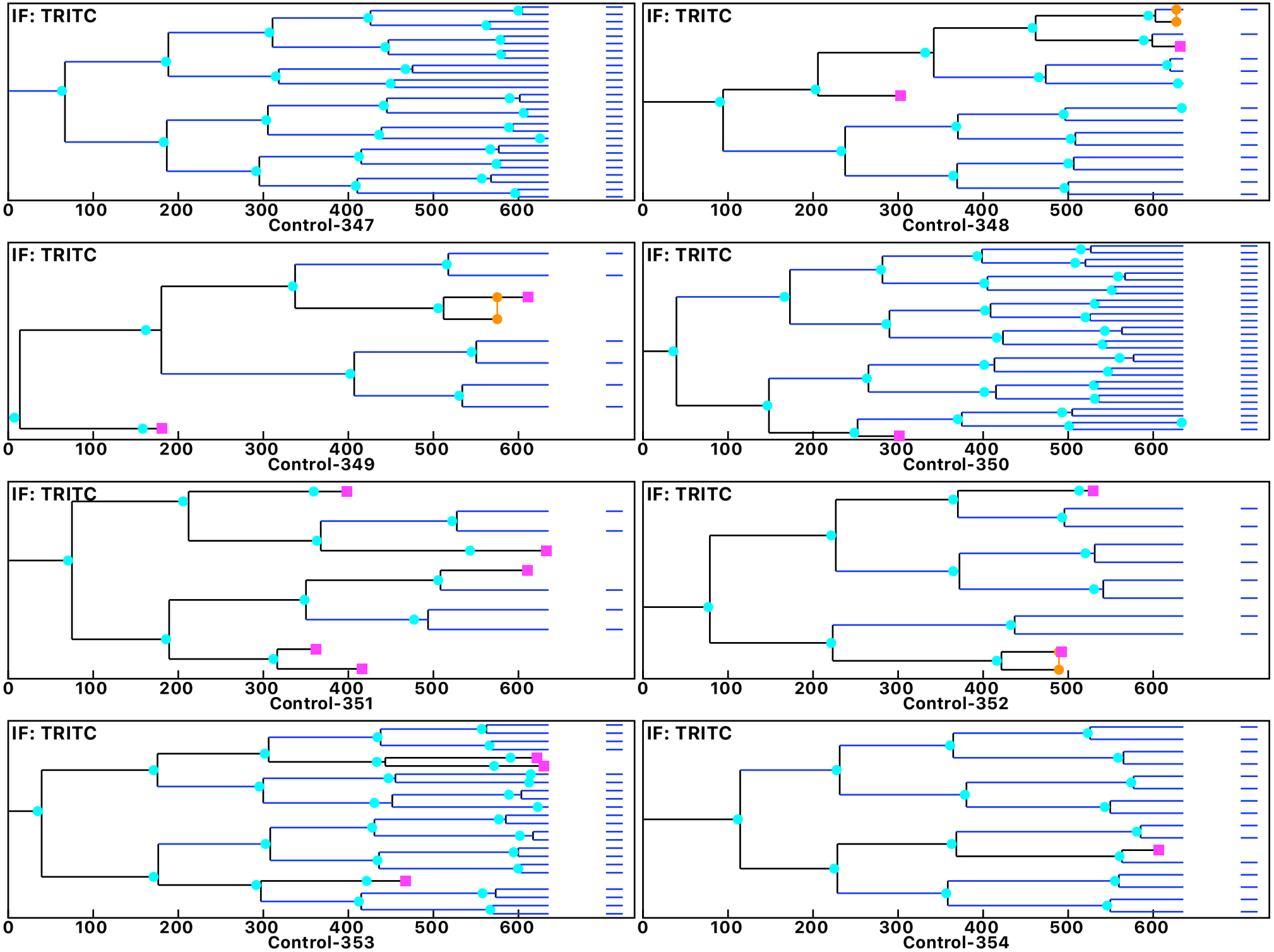




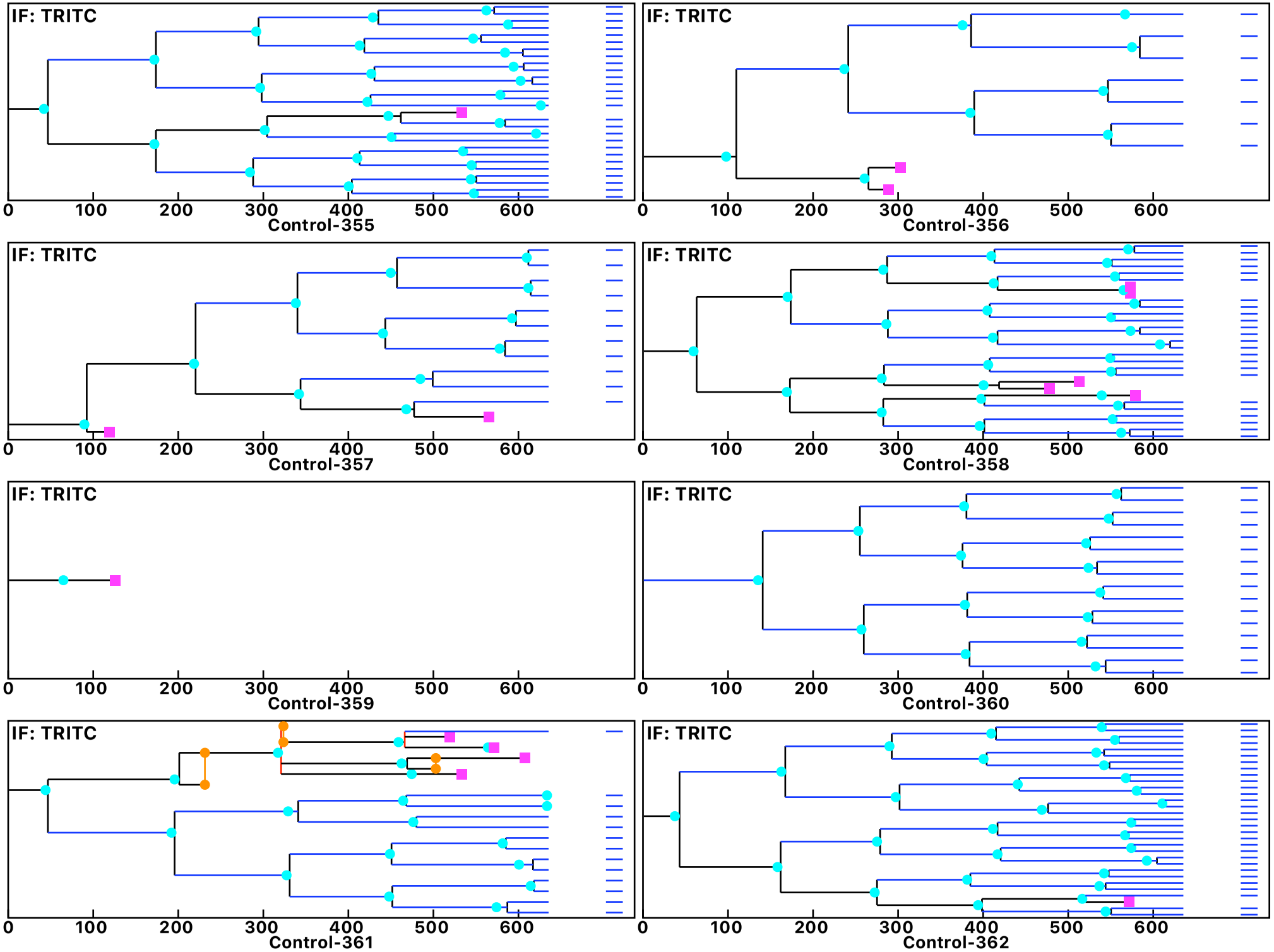
Analysis: HeLa, Treat.: Control, Cell: HeLa



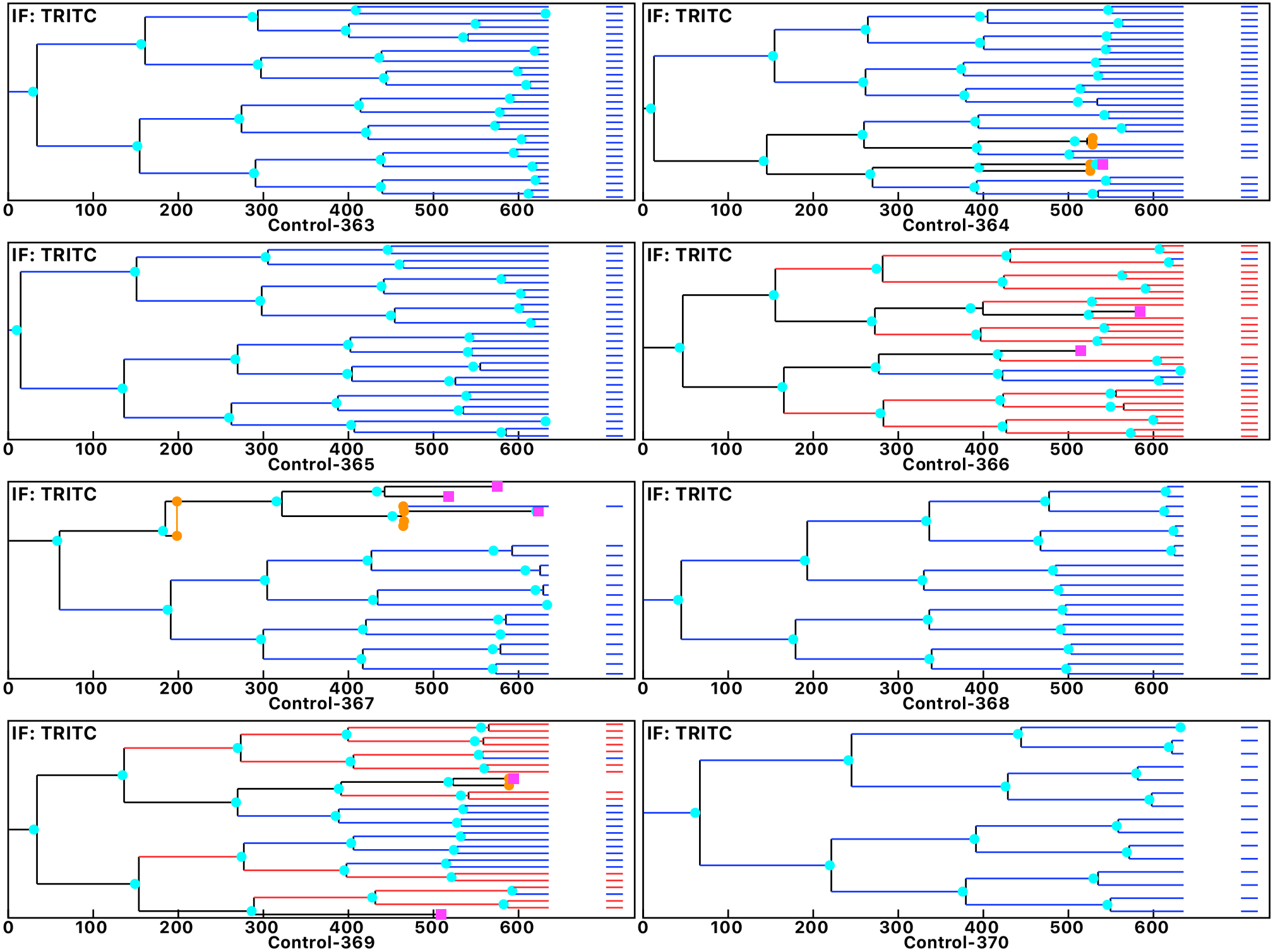
Analysis: HeLa, Treat.: Control, Cell: HeLa



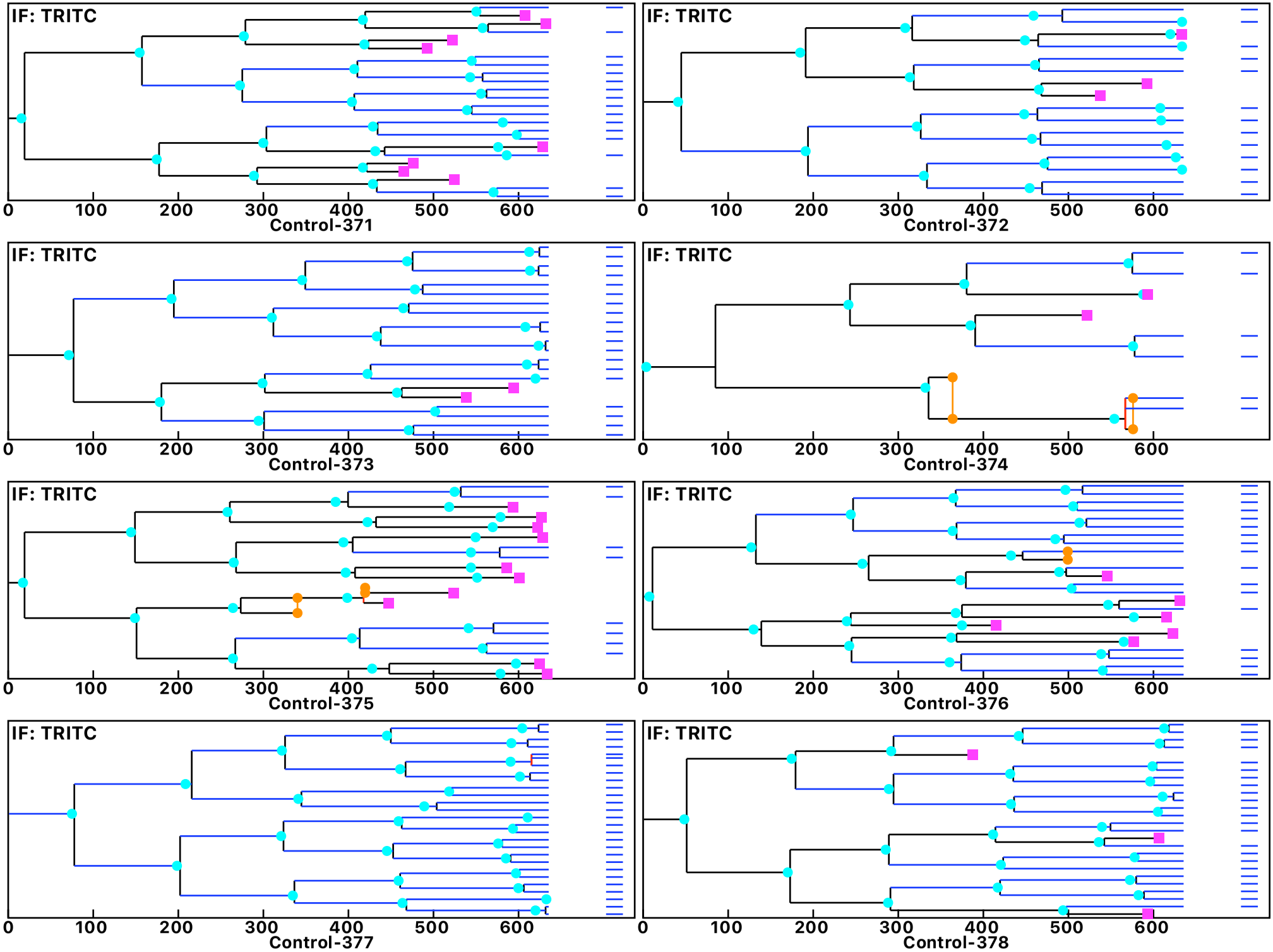
Analysis: HeLa, Treat.: Control, Cell: HeLa



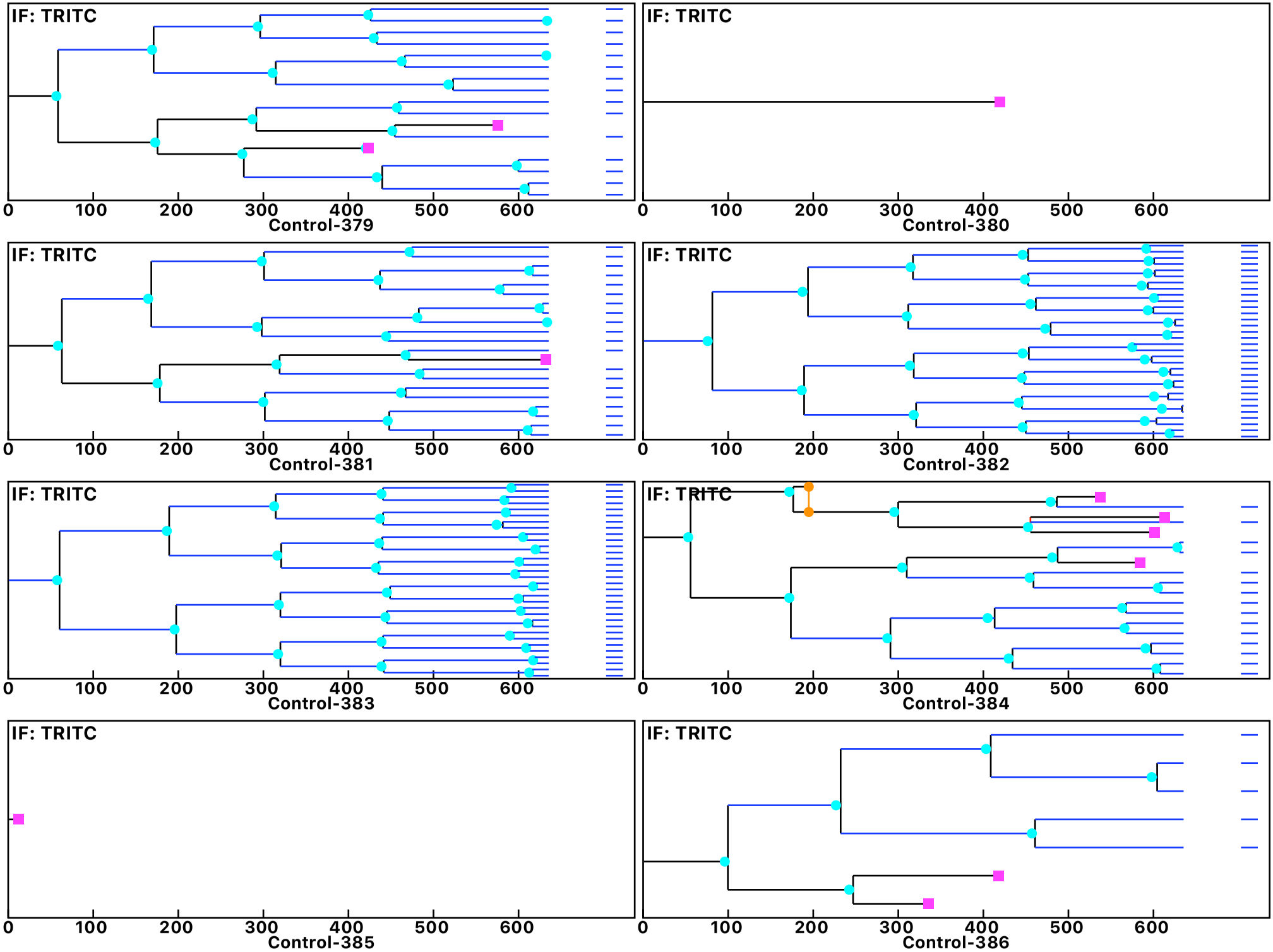
Analysis: HeLa, Treat.: Control, Cell: HeLa



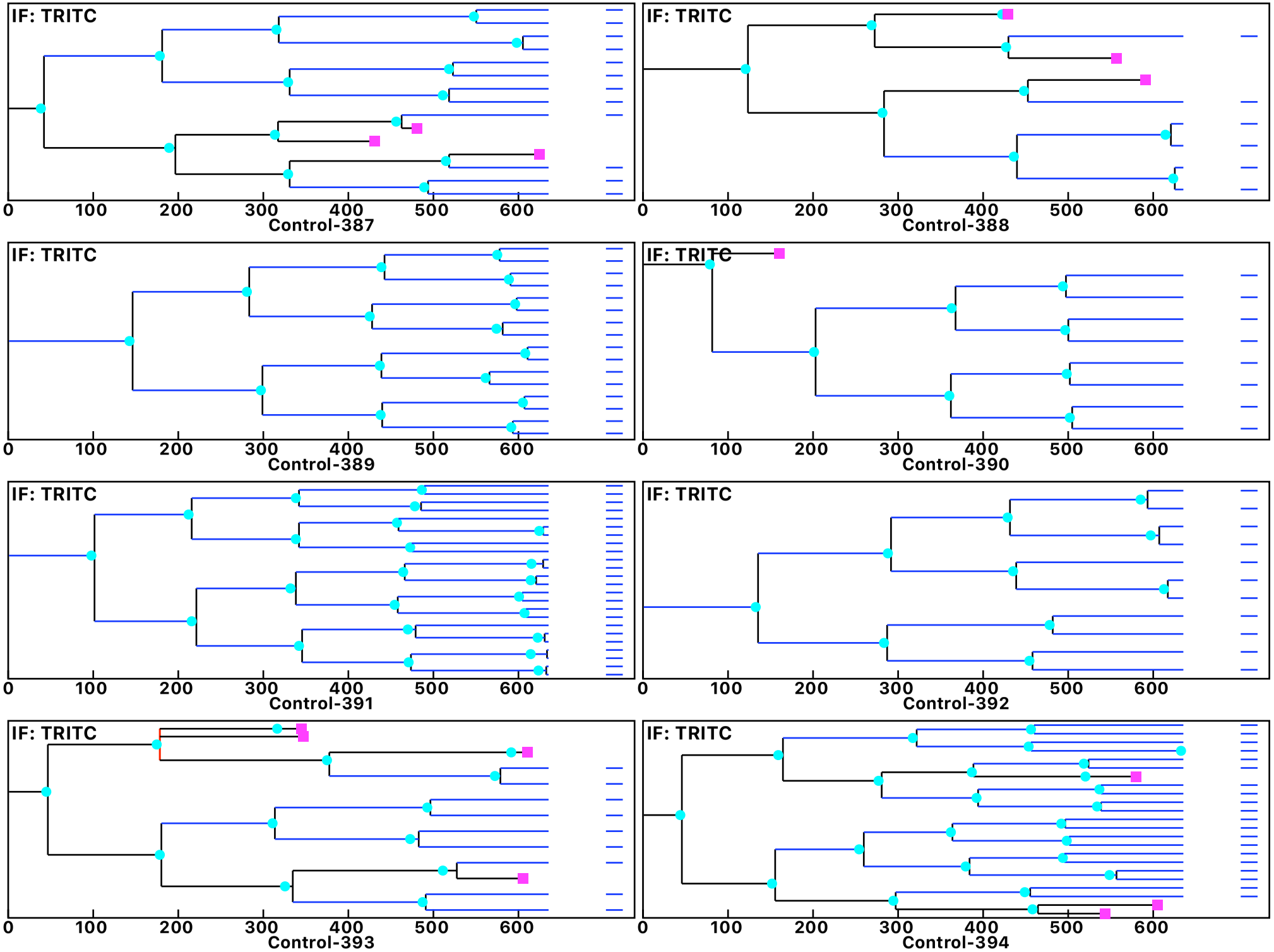
Analysis: HeLa, Treat.: Control, Cell: HeLa



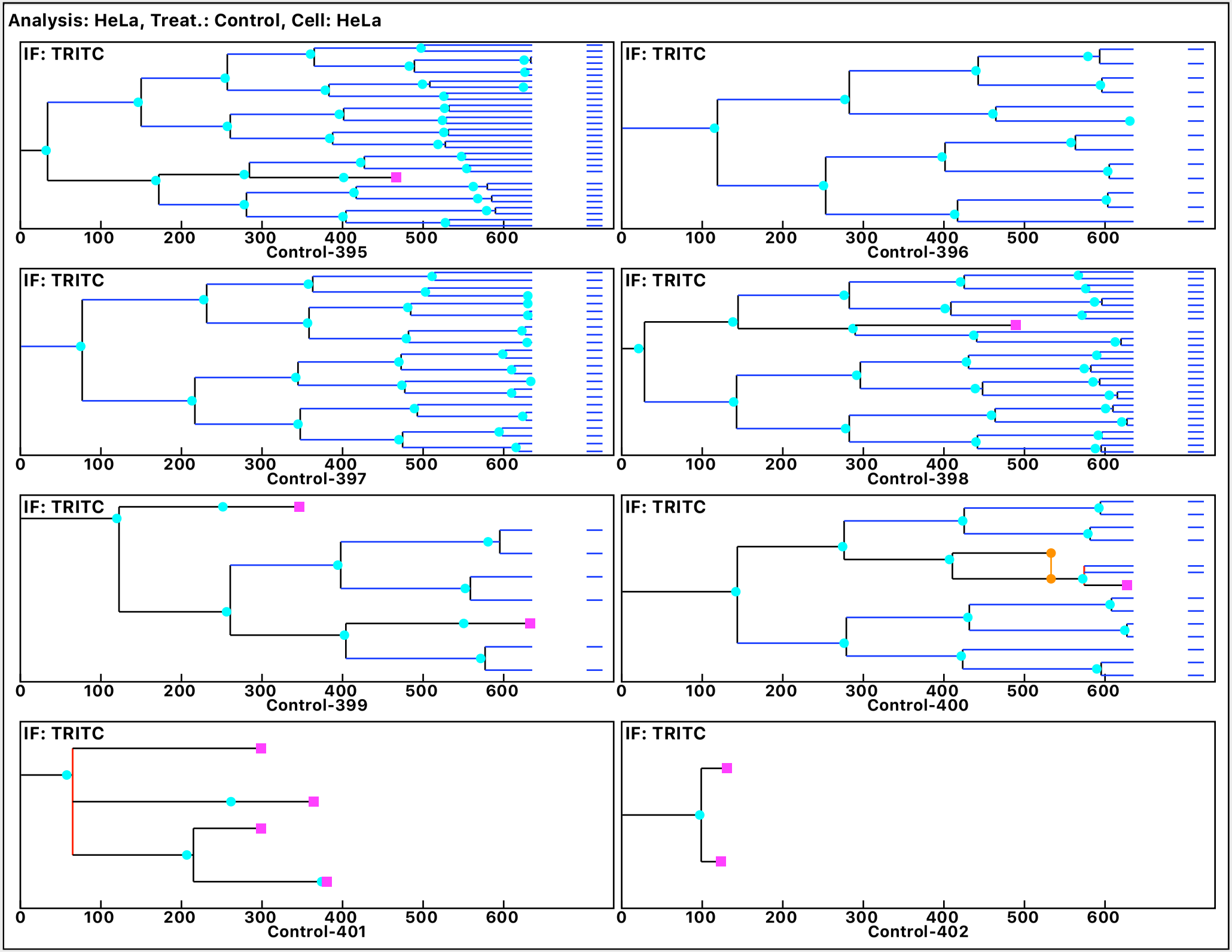
Analysis: HeLa, Treat.: Control, Cell: HeLa



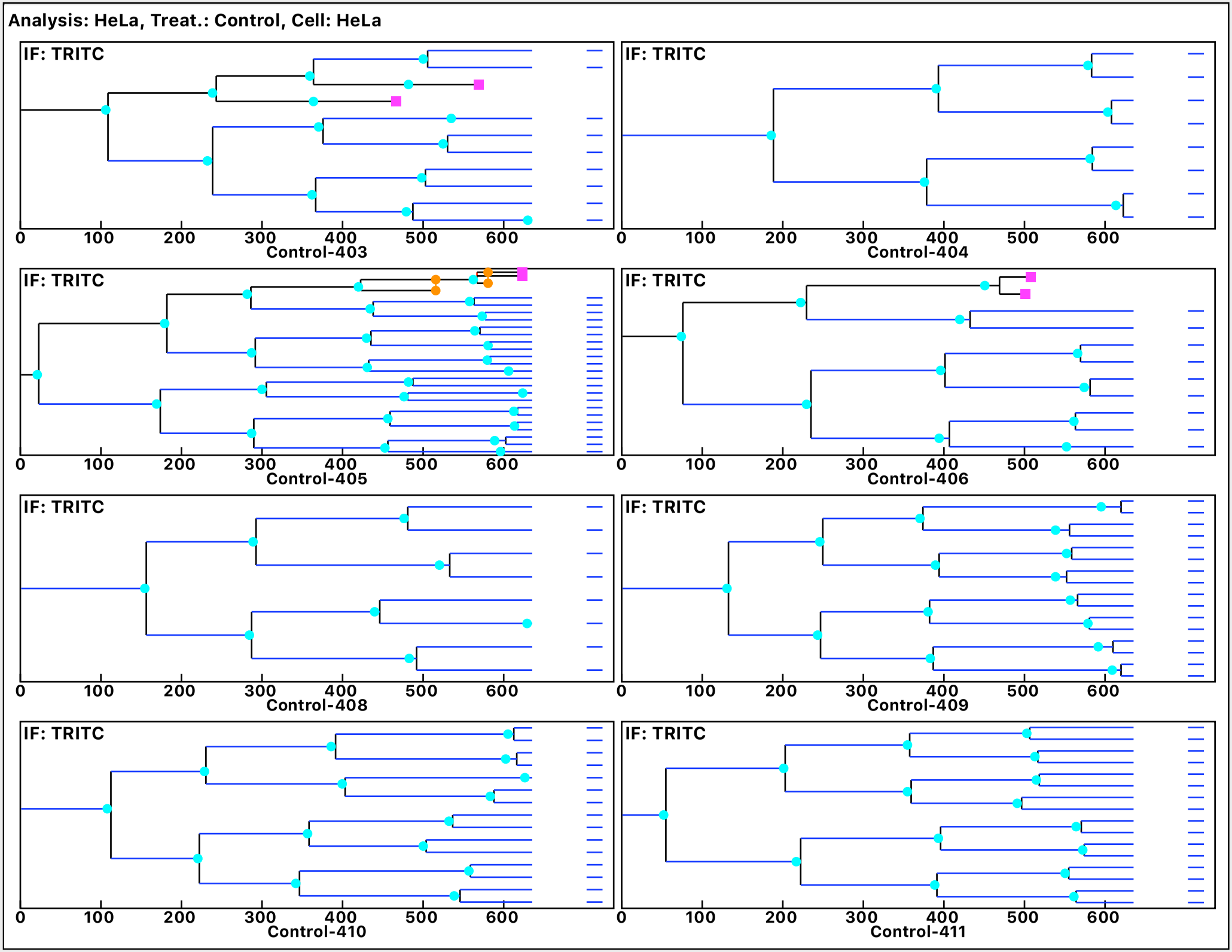
Analysis: HeLa, Treat.: Control, Cell: HeLa



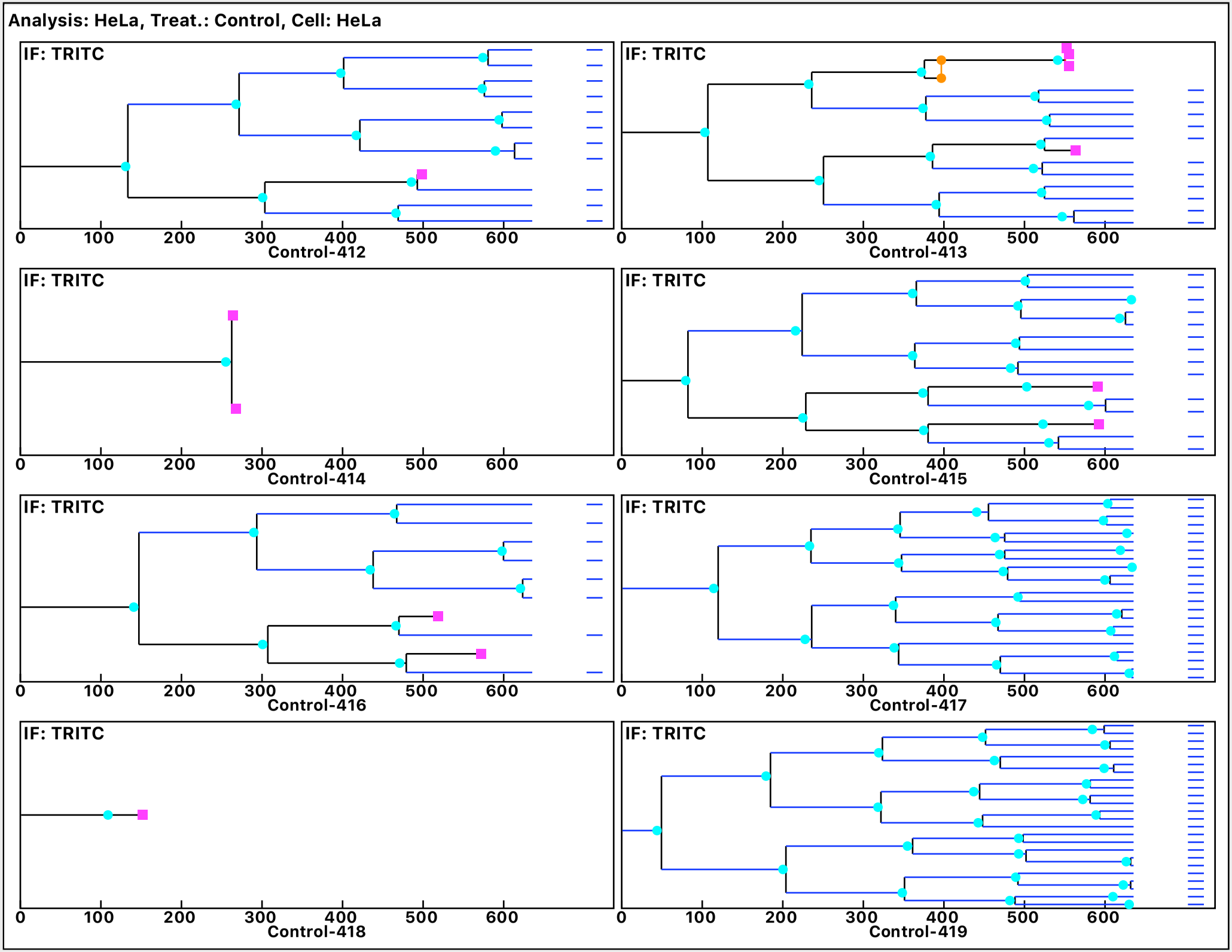
Analysis: HeLa, Treat.: Control, Cell: HeLa



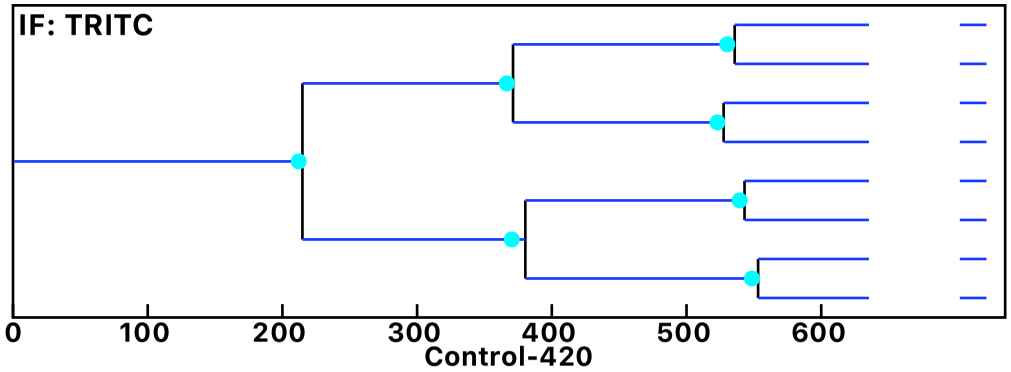
Analysis: HeLa, Treat.: Control, Cell: HeLa



Analysis: HeLa, Treat.: Control, Cell: HeLa

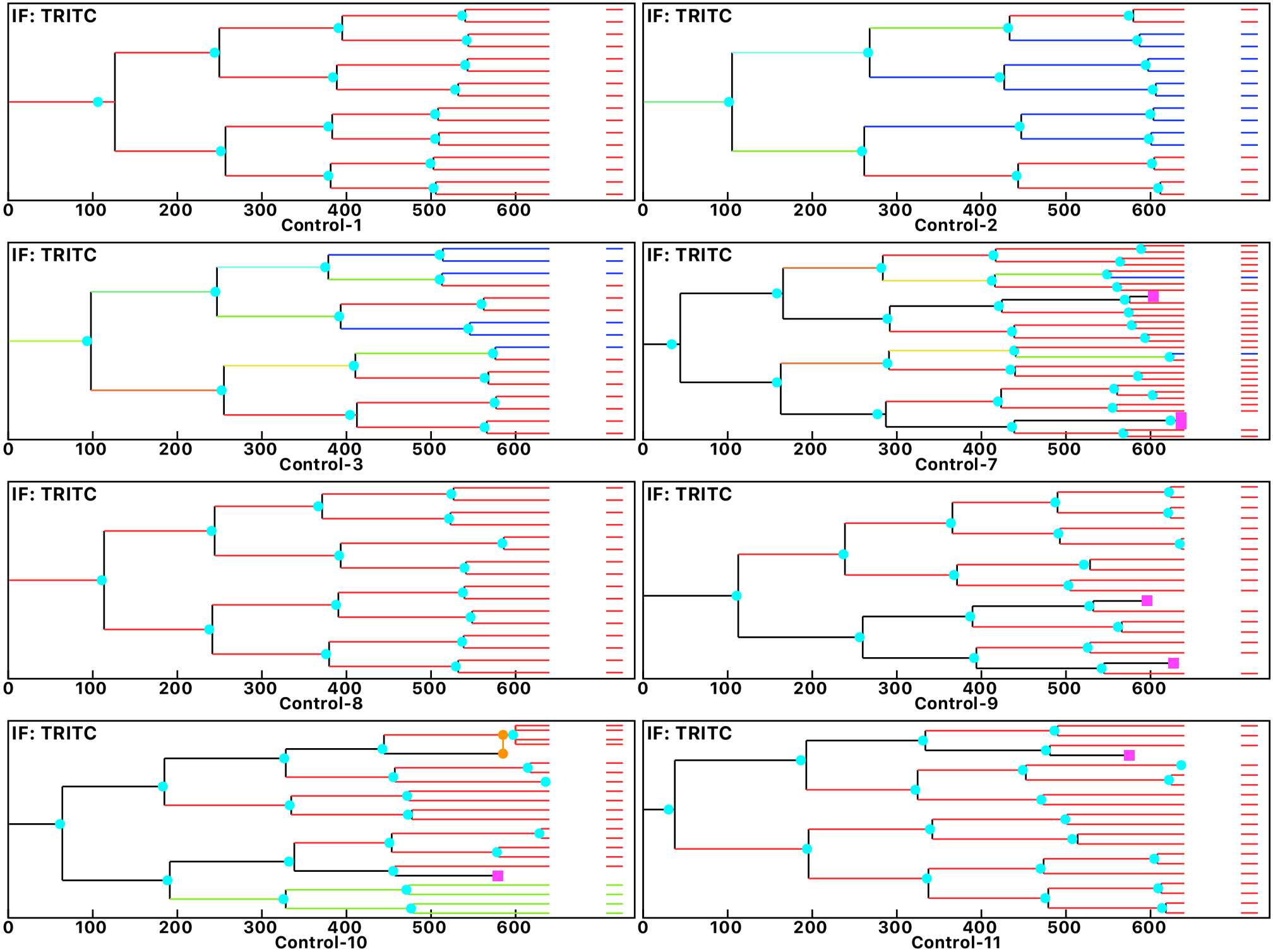


Analysis: HeLa, Treat.: Control, Cell: HeLa

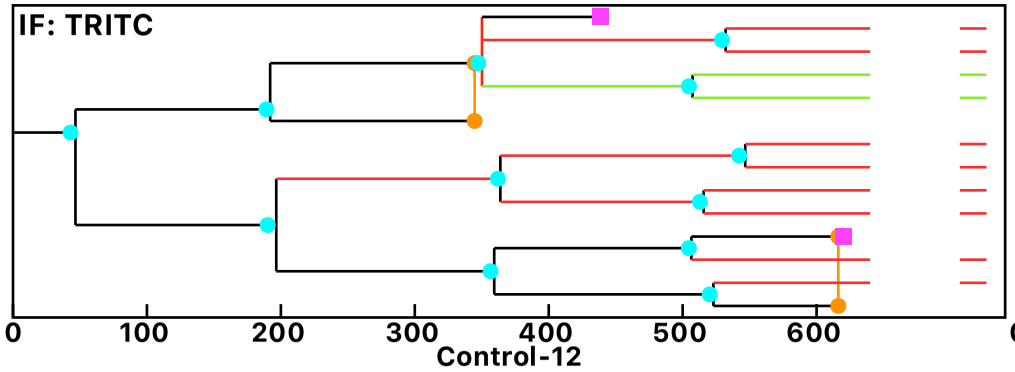


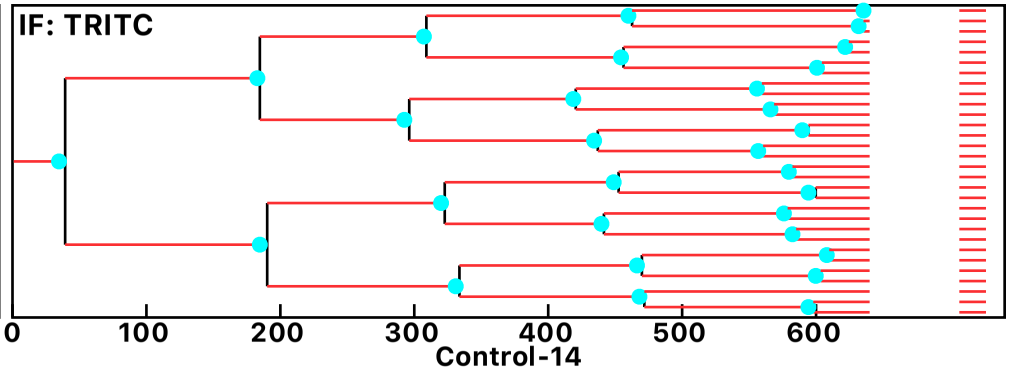
Supplementary Figure S9

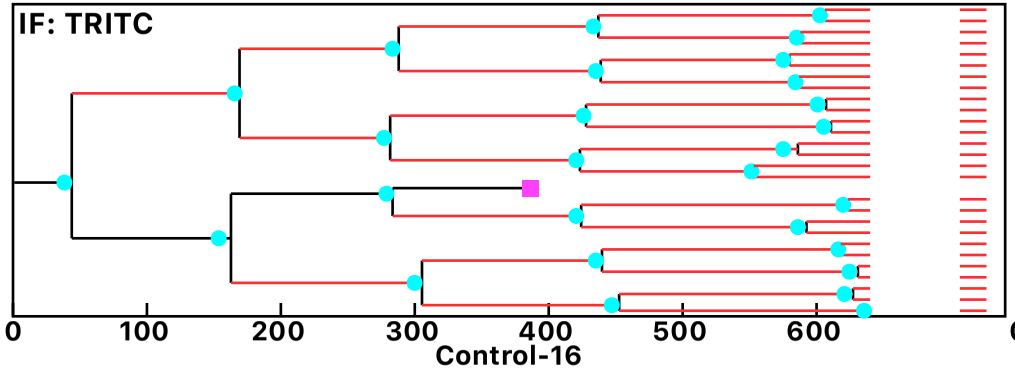
Analysis: SNA, Treat.: Control, Cell: HeLa

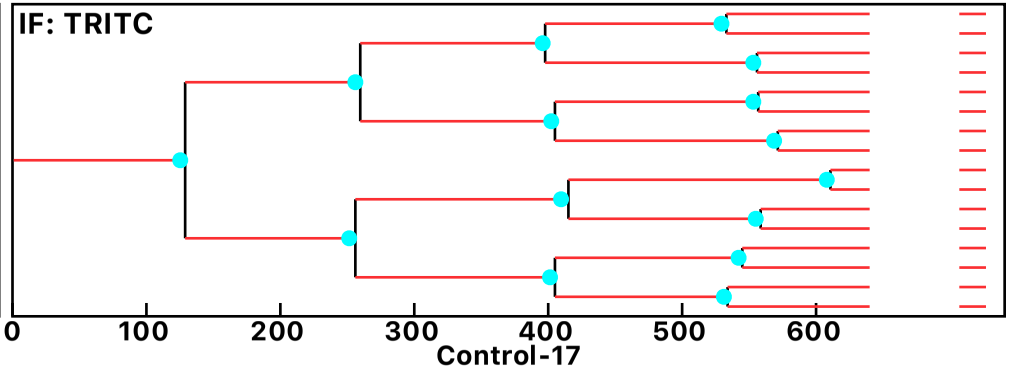


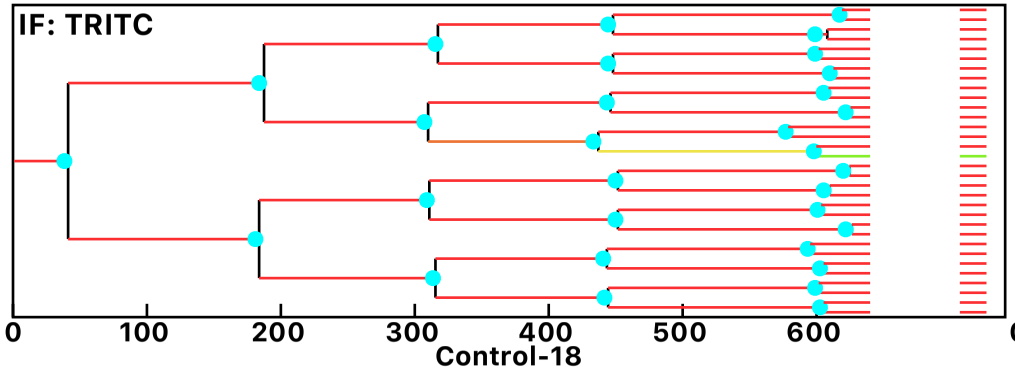
Analysis: SNA, Treat.: Control, Cell: HeLa

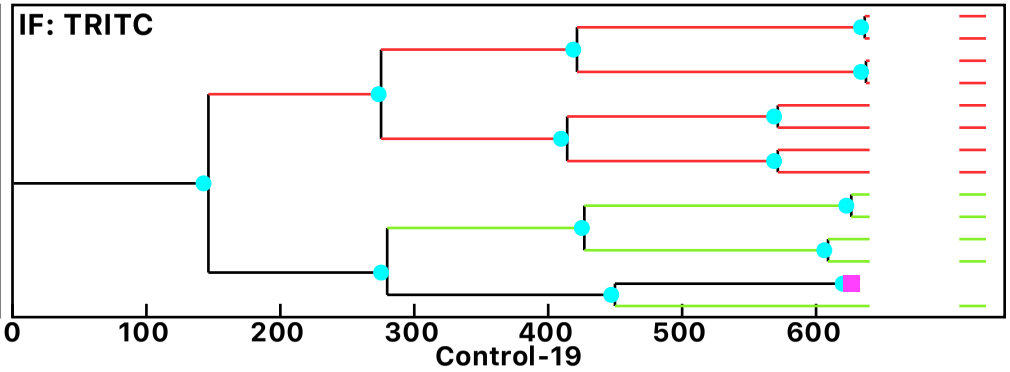
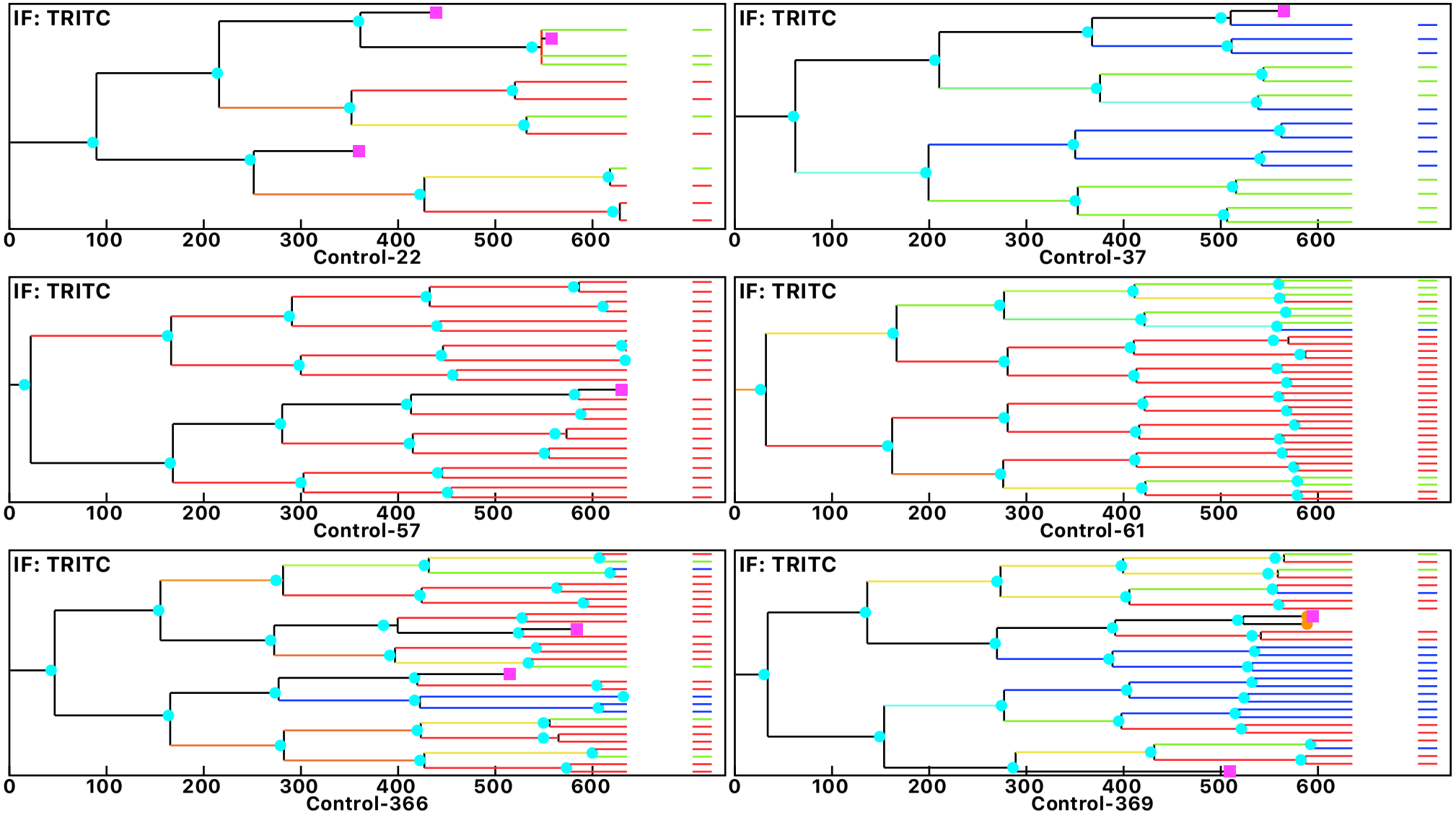












Analysis: SNA, Treat.: Control, Cell: HeLa