High throughput image analysis of trypanosomatids – differential DNA staining and colour deconvolution

Richard Wheeler, Eva Gluenz, Keith Gull

The Sir William Dunn School of Pathology, University of Oxford, Oxford, UK

Supported by funding from:

Programme grant and DPhil studentship

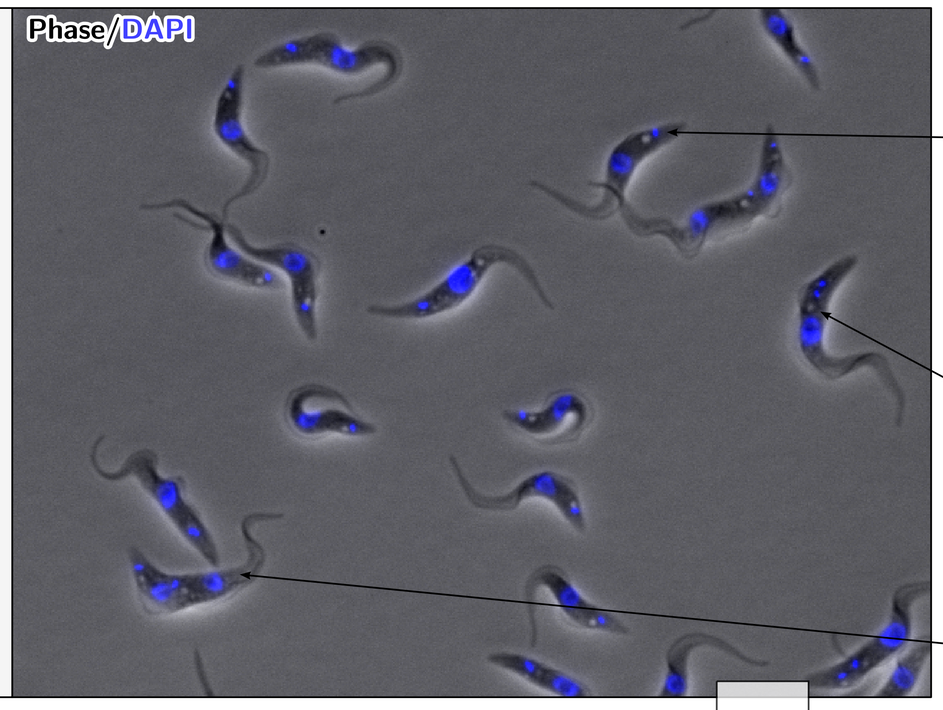
Via the Oxford Centre for Integrative Systems Biology

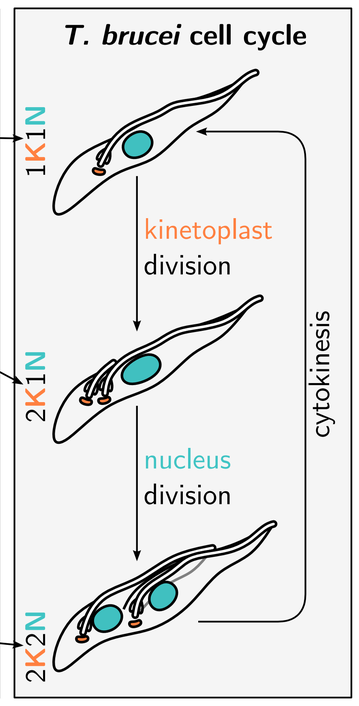
1. DNA in trypanosomatids

Kinetoplastids are an order of parasitic protozoa which includes the major human pathogens *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania* spp. Kinetoplastids share several distinctive features:

- A highly structured cell body with a microtubule-based cytoskeleton and a single flagellum
- A kinetoplast; the mitochondrial DNA which is found at the base of the flagellum
- Single copies of vital organelles including the nucleus, kinetoplast, flagellum and mitochondrion

Posession of two large DNA-containing organelles complicates automated image analysis; the kinetoplasts and nuclei must be reliably identified.





Trypanosomatids have a complex and rigid ultrastructure and a highly choreographed cell cycle. Detection of deviation from normal morphology and the normal cell cycle is used to:

- Analyse reverse genetic mutants to infer protein function
- Screen potential drug candidates to detect anti-trypanosomatid activity
- Detect and analyse life-cycle stage differentiation, which normally involves modified division

Cell morphology and behaviour of cells through the cell cycle is routinely analysed by the nucleus and kinetoplast number. This is the analysis technique I have automated *and* extended.

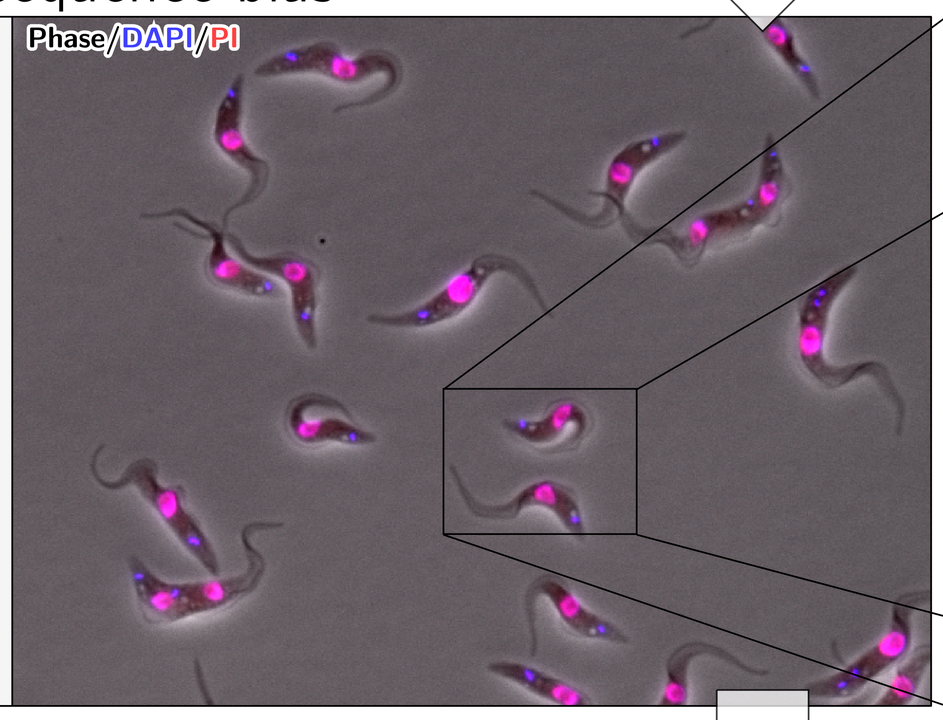
2. Identifying DNA by sequence bias

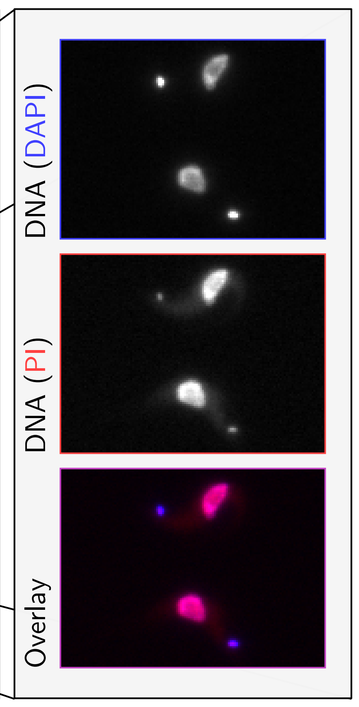
Using two different DNA stains we can identify the kinetoplast and nucleus automatically; different DNA stains have different sequence binding biases so label kinetoplasts and nuclei differently.

There are two main classes of DNA binding small molecule fluorescent stains:

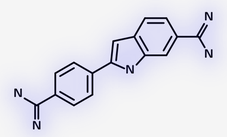
- Base pair intercalating
  e.g. propidium iodide (PI), SYBR green, ethidium bromide
- Minor groove binding
  e.g. DAPI, Hoechst

Minor groove binding stains tend to prefer A-T rich DNA sequences. Kinetoplast DNA is A-T rich so binds DAPI strongly.

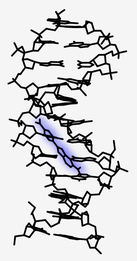




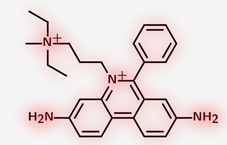
DAPI



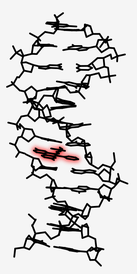
- Binds the DNA minor groove
- A-T sequence selective
- Blue fluorescence



Propidium iodide (PI)



- Intercalates DNA base pairs
- Low sequence specificity
- Red fluorescence
- Binds RNA (RNase treatment required)

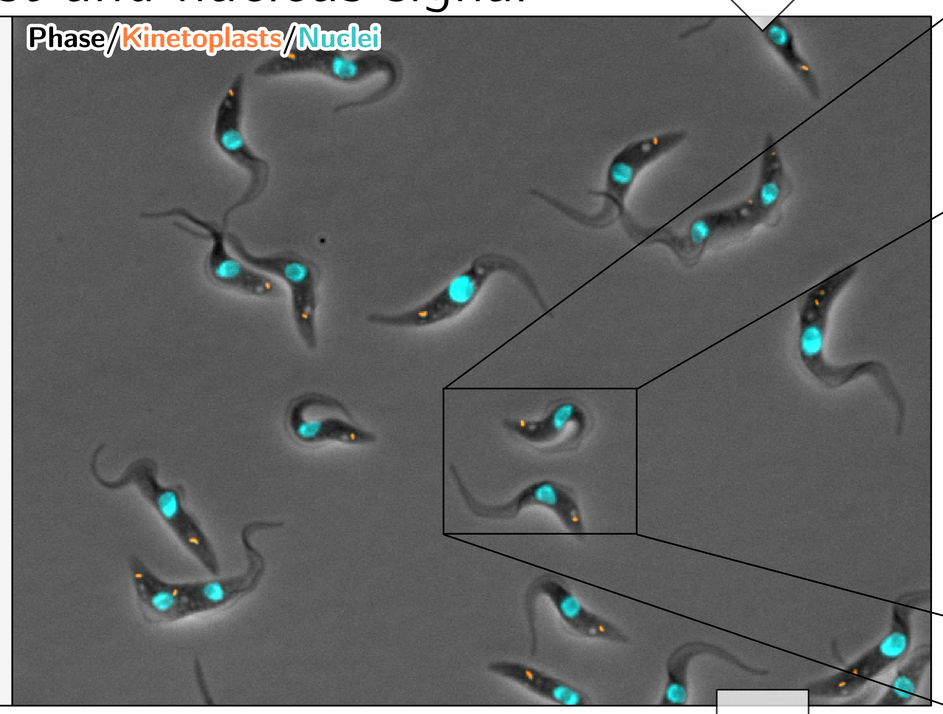


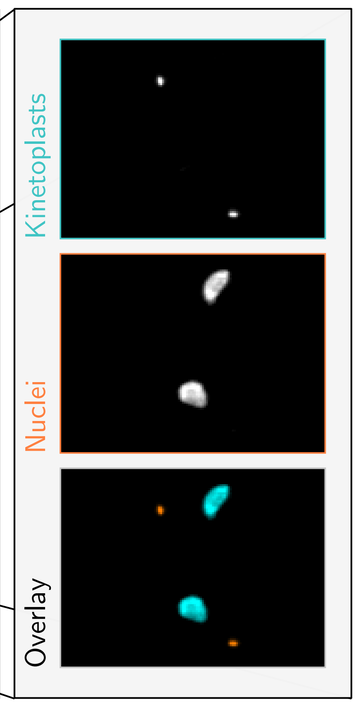
3. Separating kinetoplast and nucleus signal

Using colour deconvolution we can separate the kinetoplasts and nuclei from the DAPI and PI fluorescence images to to simplify automatic analysis.

Colour deconvolution is calculated using reference intensity values and the DAPI and PI values of a pixel. Image processing takes four steps, all are fully automatic:

1. Measure chromatic aberration between the DAPI and PI images
2. Correct chromatic aberration with a linear scaling of the PI image
3. Measure the reference DAPI and PI values for kinetoplasts and nuclei
4. Perform colour deconvolution to produce kinetoplast and nucleus images





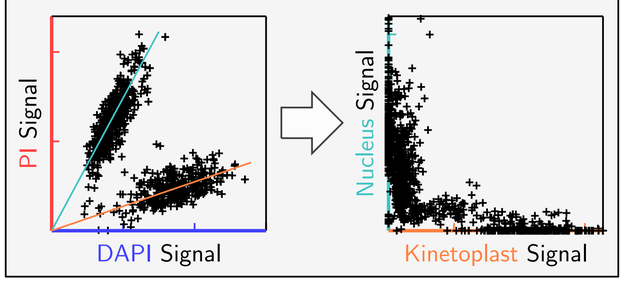
Mathematically colour deconvolution is a change in basis in two dimensions.

$$\mathbf{p'}=M^{-1}\mathbf{p}$$

Where **p** is the DAPI/PI pixel value and **p'** is the kinetoplasts/nuclei pixel value. The transformation matrix, M, is given by:

$$M=(\mathbf{k},\mathbf{n})$$

Where **k** and **n** are the reference DAPI/PI values for a kinetoplast and nucleus respectively.

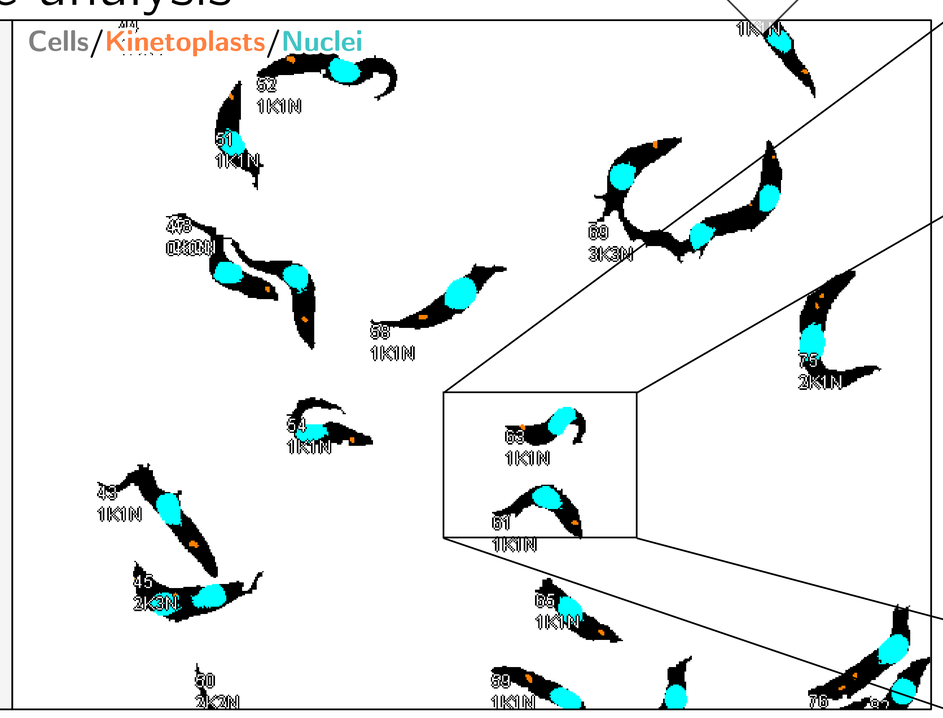


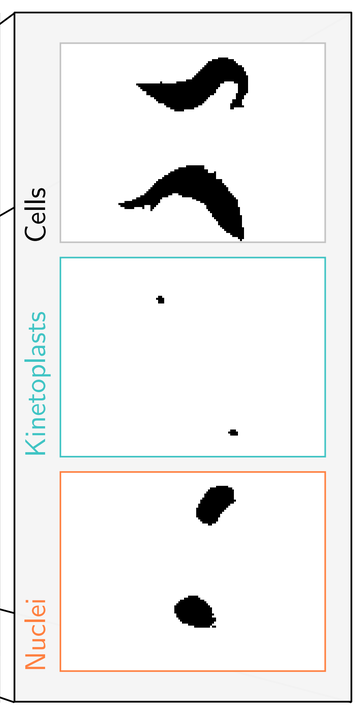
4. Automation of image analysis

Colour deconvolution generates quantitative and easy to analyse images of the kinetoplasts and nuclei. We have designed tools to automate analysis of these images.

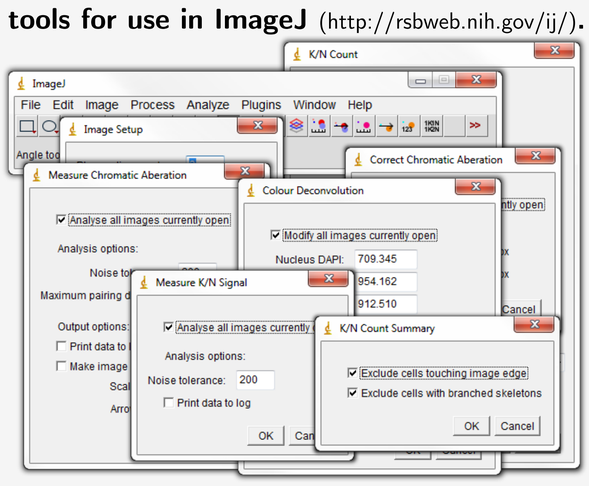
Image analysis uses simple intensity thresholding to identify cells, nuclei and kinetoplasts. Automated analysis takes thres steps, all are fully automatic:

1. Intensity threshold kinetoplast and nucleus images, subtract background and threshold phase image
2. Analyse cell shape using a medial axis transform and extract cell width, length and shape data
3. Measure kinetoplast and nucleus properties from the deconvolved images including DNA quantity and organelle location





All these processes have been fully automated in the form of a set of macro tools for use in ImageJ (http://rsbweb.nih.gov/ij/).



A typical image can be processed in 10–15s. With around 50–100 cells per image these macros can analyse around 20000 cells per hour.