An image based high throughput screen to identify regulators of Imp containing RNP granules

Fabienne de Graeve, Somia Rahmoun, Djampa Kozlowski, Nicolas Cedilnik, Eric Debreuve, Xavier Descombes, Florence Besse

An image based high throughput screen to identify regulators of Imp containing RNP granules

Fabienne De Graeve[1], Somia Rahmoun[2], Djampa Kozlowski[2], Nicolas Cedilnik[2], Eric Debreuve[3,4], Xavier Descombes[2,4], Florence Besse[1,4].

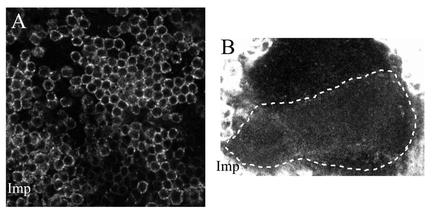
[1] UCA, UNS, iBV CNRS UMR 7277, University of Nice Sophia-Antipolis, Parc Valrose, 06108 Nice Cedex 2, France
[2] UCA, INRIA, I3S 2000, route des Lucioles - Les Algorithmes - bât. Euclide B - BP 121, 06903 Sophia Antipolis Cedex, France
[3] UCA, I3S, INRIA 2000, route des Lucioles - Les Algorithmes - bât. Euclide B - BP 121, 06903 Sophia Antipolis Cedex, France [4] Morpheme team members

I. Context

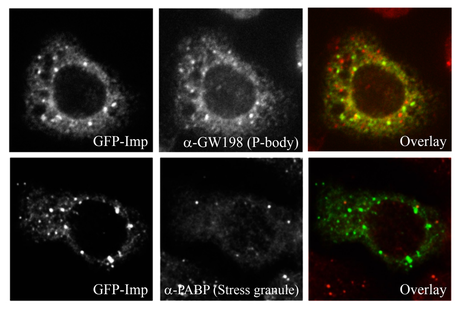
In vivo, RNAs and proteins are frequently packaged into diverse dynamic macromolecular structures named mRNP granules. These assemblies form upon phase separation of individual RNA and protein components, a process involving the establishment of multivalent weak interactions and their regulations via post-translational modifications. Defects in their properties have been associated with several human pathologies. However, our knowledge of these dynamic structures relies essentially on the study of P bodies and stress granules.

We are interested in the highly conserved RNA binding protein Imp whose mammalian counterpart's overexpression correlates with poor prognosis in several cancers. *In vivo*, Imp is present in cytoplasmic RNP granules, distinct from P-bodies and visible both in neuronal cell bodies and axons. They are also detected in Drosophila S2R^{+} cultured cells.

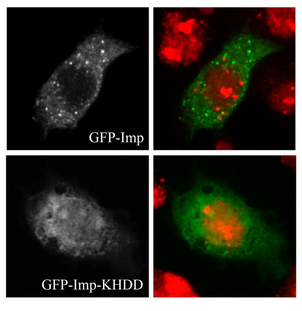
Taking advantage of this cellular model, we have undertaken a genome-wide RNAi-based visual screen to identify factors that regulate the properties of Imp-containing granules. This implies combining high throughput microscopy with the development of a computational pipeline for automatic image analysis. This pipeline first segments and discriminates healthy from dead nuclei, storing this information in an interactive SQLite database that enables experimental quality control. Then, GFP-Imp granules are detected using the SPADE algorithm in the cytoplasm of healthy cells. Data from the pilot screen we have performed to validate the experimental design and develop our pipeline for data mining are presented.



In vivo, Imp is present in RNP granules, both in neuronal cell bodies (A) and axons (B). Shown here, the mushroom body neurons. (courtesy of Kavya Pushpalatha)

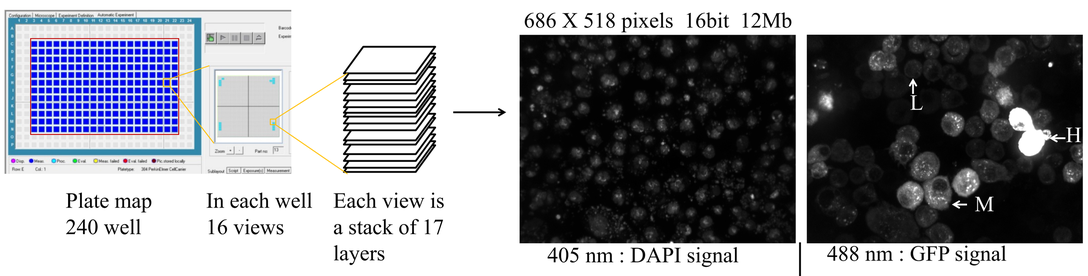


In S2R+ cells Imp granules are different from P bodies and stress granules



RNA binding activity is required for the assembly of Imp granules *in vitro* and *in vivo* (not shown)

II. Pilot screen



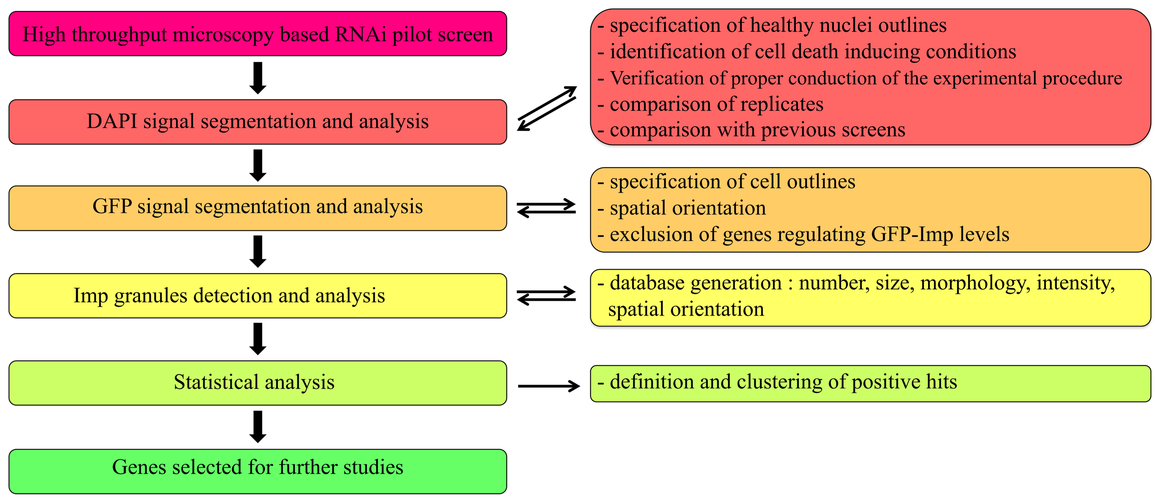
The pilot screen was performed at the DRSC (Boston, USA)

Per well 20.000 cells were transiently transfected (L : low, M : medium and H : high level of GFP-Imp expression)

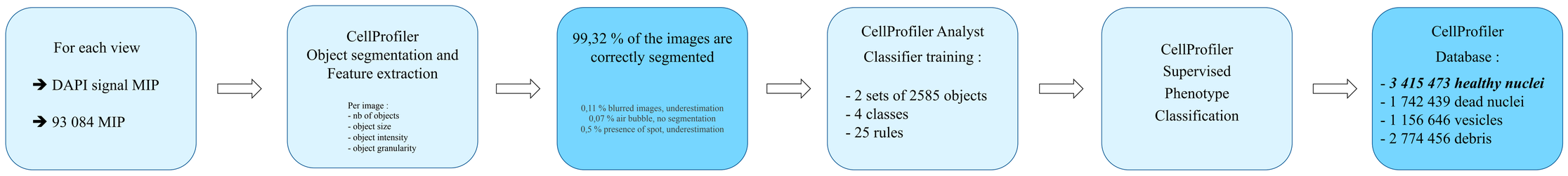
Two sublibraries were screened :
Phosphatases – Kinases sublibrary : 563 genes, 8 plates
RNA Binding Proteins sublibrary : 406 genes, 4 plates

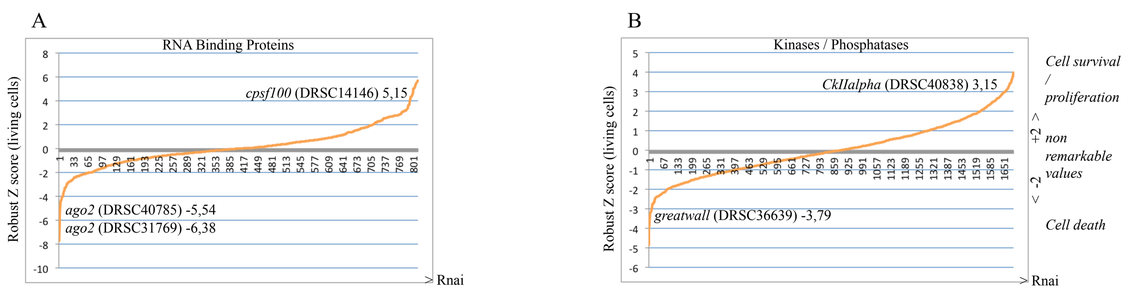
Experiments done in replicates :
24 plates*240 wells*16 views*17 stacks*2 channels => 3.10^6 images !

III. Workflow



IV. Experimental procedure quality check and selection of healthy cells





Distribution of the robust z-scores from the pilot screen using living cells as readout (we used median normalisation per plate (excluding controls)). In this pilot screen, 812 different RNA strands were used to reduce the expression of one of the 406 genes encoding a RNA binding protein (A). 1696 different RNA strands were used to reduce the expression of one of the 563 genes encoding a kinase or protein phosphatase (B). Similarly to Mohr et al. we found that *ago2* and *cpsf100 rnai* downregulation favor cell death and cell survival respectively. These results validate the reproducibility of the experimental procedure and our statistical approach.

Visual inspection of raw data indicates :

(1) The cell population inside each well is heterogeneous.

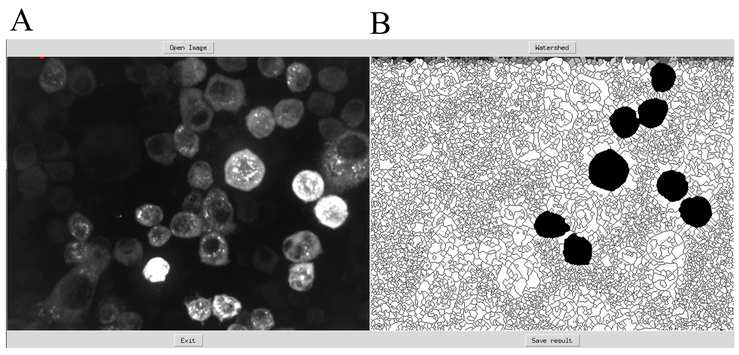
(2) In transfected cells, expression levels of the GFP-Imp molecule are variable : low, medium or high. Only cell expressing GFP-Imp at a low level are interesting.

(3) Depending on the cells, the RNP-Imp granules may be more or less numerous, larger or smaller, they may accumulate locally or be distributed homogeneously throughout the cytoplasm.

(4) Inactivation of most of the genes tested does not induce a detectable phenotype by eye

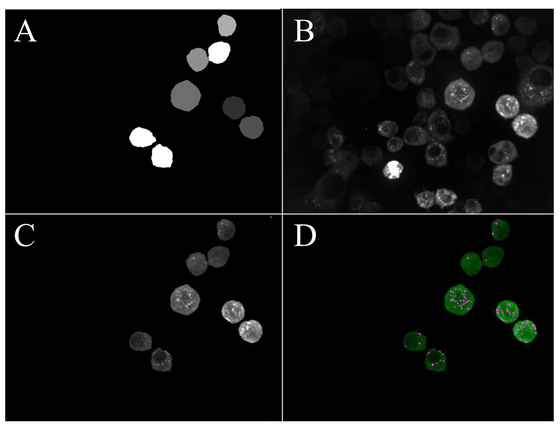
(5) Hits with more granules were detected (about 1%)

V. Semi-Automatic Cell Segmentation : from GFP MIP images to cell masks



GFP MIP images (A) are first locally-normalized to reduce the effect on a non-uniform illumination. Then, a watershed-based segmentation produces the corresponding oversegmented (superpixels) image (B). Finally, the interface allows to manually select the superpixels to segment the cells of interest by creating masks (B). In the future, an automatic process for superpixels merging will be developed.

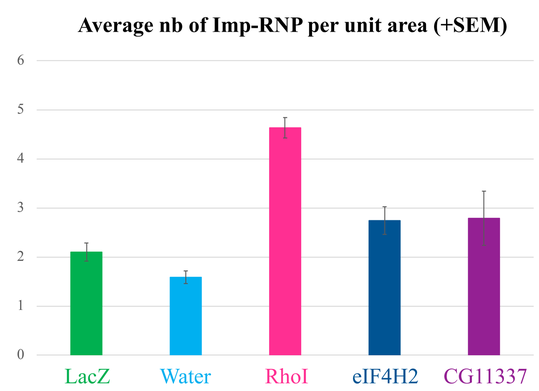
VI. SPADE : an algorithm for small particle detection



The mask generated by the supervised segmentation algorithm (A) is used to extract from the original MIP image pretreated for enhanced contrast normalization (B), the GFP positive cytoplasm of healthy cells whose intensity ranges from 100 to 1000 (C). This latter is analyzed by SPADE that combines the point process concept of minimizing an energy function with the use of a pre-defined dictionary of shapes, allowing the detection of small objects, ranging from 4 to 25 pixels in a heterogeneous context (D). 1. http://spade.gforge.inria.fr/ 2. https://pypi.python.org/

VII. Preliminary data :

Plate_Well	RNAi	(nb of RNP per pixel)*1000
461T1_C03	LacZ	1,89
461T1_E21	LacZ	2,65
461T1_F15	LacZ	1,41
461T1_G05	LacZ	1,63
461T1_I08	LacZ	2,16
462T2_C03	LacZ	2,54
461T1_C16	H2O	1,60
461T1_D06	H2O	1,65
461T1_D22	H2O	1,55
461T1_H03	H2O	1,10
462T2_C16	H2O	1,74
461T1_C11	RhoI	4,57
461T1_G13	RhoI	4,25
461T1_I04	RhoI	5,03
461T1_K06	RhoI	6,06
462T2_C11	RhoI	3,53
461T1_F14	eIF4H2	1,92
461T3_F14	eIF4H2	3,46
463T1_K07	eIF4H2	2,87
461T1_M11	CG_11337	2,79



Images from different control wells (LacZ and water) of one or two independent plates as well as images from 3 possible hits were analysed as described before (IV to VI). Results show robustness of the controls. As for quality control, we will use median normalisation per plate (excluding controls). Hits will be identified on the basis of their average number of IMP containing RNP per unit area as well as other criteria to be tested.

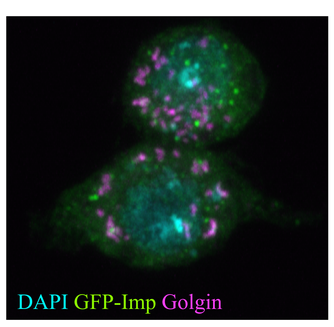
VIII. Perspectives

Automate the merging process for image treatment

Develop the statistical framework to detect hits

In vivo functional analysis of hits

Improve the experimental readout using additional markers (phalloïdin, organelle markers)



Collaborators

DRSC (Harvard, USA)
- *high throughput RNAi screen*

Fundings