PMLs and SPPs

David A. Stephens

Department of Mathematics
Imperial College London

d.stephens@imperial.ac.uk
http://stats.ma.ic.ac.uk/~das01/

21st February 2006

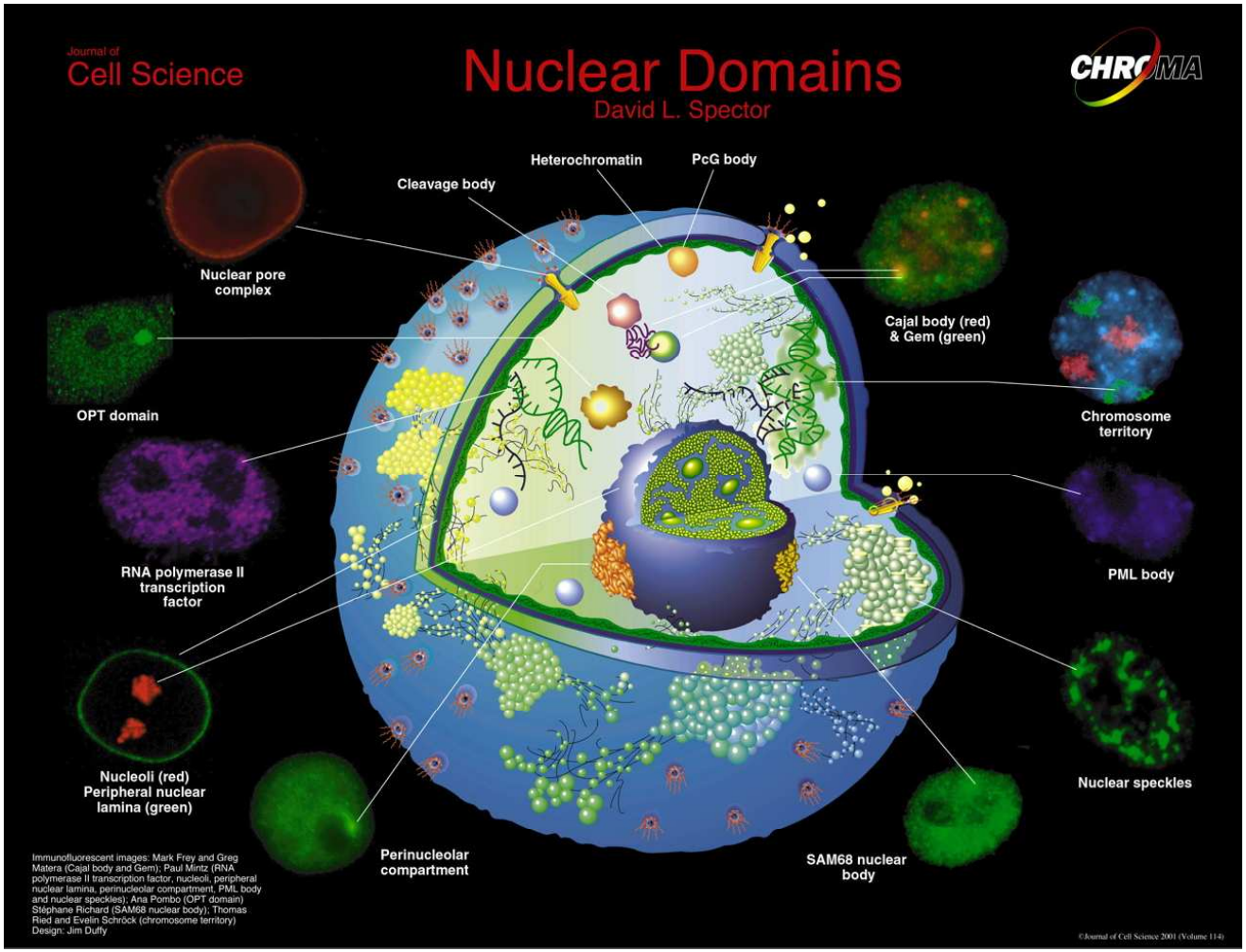
Biological Objective

To understand organization in the mammalian cell nucleus via analysis of its components.

> *The mammalian cell nucleus is a membrane-bound organelle that contains the machinery essential for gene expression. Although early studies suggested that little organization exists within this compartment, more contemporary studies have identified an increasing number of specialized domains or subnuclear organelles within the nucleus .... An extensive effort is currently underway by numerous laboratories to determine the biological function(s) associated with each domain.*

Spector (2001) J. Cell Sci.

Nuclear Domains



Spector (2001) J. Cell Sci.

Nuclear Domains

- Cajal body
- Nuclear Gems
- Chromatin proteins
- PML body
- Splicing Factor Enriched Speckles
- Sam68 body
- Nuclear Diffuse proteins
- Perinuclear Compartment
- Nucleolar proteins
- Nuclear Lamina
- RNA pol II transcription factors
- OPT domain
- Nuclear Pore
- Cleavage body
- Heterochromatin proteins
- Paraspeckles

Nuclear Domains

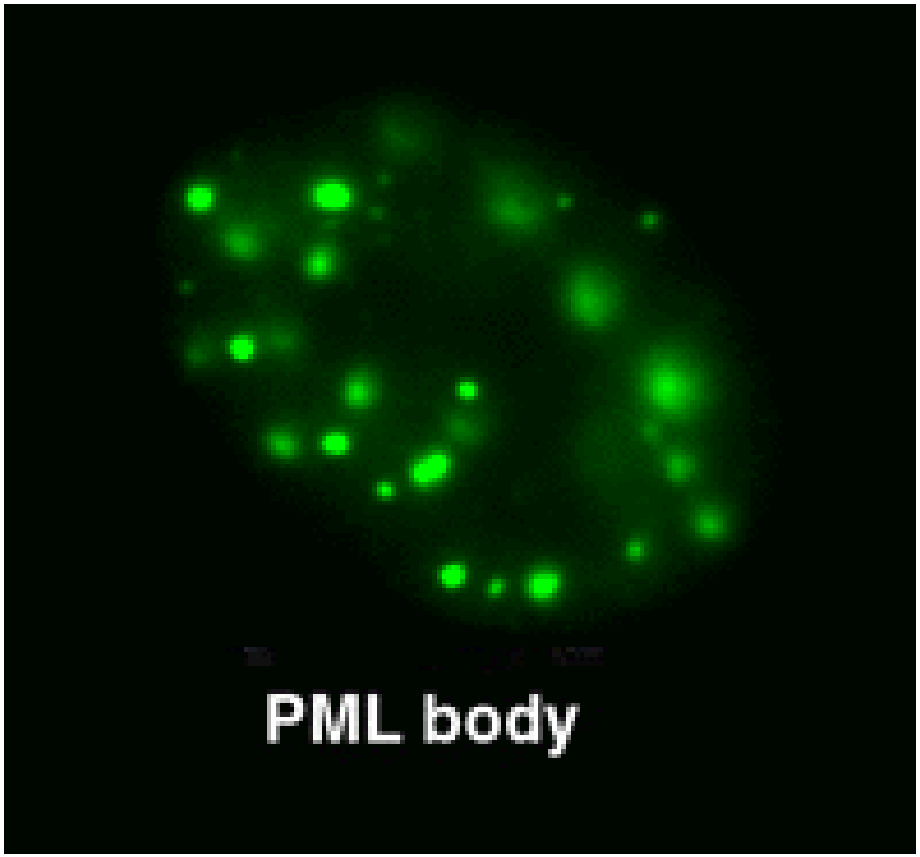
- Cajal body
- Nuclear Gems
- Chromatin proteins
- PML body
- Splicing Factor Enriched Speckles
- Sam68 body
- Nuclear Diffuse proteins
- Perinuclear Compartment
- Nucleolar proteins
- Nuclear Lamina
- RNA pol II transcription factors
- OPT domain
- Nuclear Pore
- Cleavage body
- Heterochromatin proteins
- Paraspeckles

PML Bodies

The promyelocytic leukemia (PML) nuclear body is a nuclear matrix-associated structure 250-500 nm in diameter that is present in the nucleus of most cell lines ... this subnuclear domain has been reported to be rich in RNA and a site of nascent RNA synthesis, implicating its direct involvement in the regulation of gene expression ... electron spectroscopic imaging (ESI) demonstrates that the core of the PML nuclear body is a dense, protein-based structure ... which does not contain detectable nucleic acid.

Boisvert et al. (2000) J. Cell Sci.

PML Bodies



PML Bodies

- aggregation of PML and other proteins.
- vary in size from 0.3 μm to 1.0 μm in diameter.
- typical mammalian nucleus contains 10-30 of these structures.
- also called ND10, PODs (PML oncogenic domains) and Kr bodies.
- several other proteins (Sp100, SUMO1, HAUSP and CBP) have been localized to this nuclear domain.
- PML bodies have been suggested to play a role in aspects of transcriptional regulation and appear to be targets of viral infection.

Protein composition of the PML body

Protein	Function
BLM	DNA helicase; complexes with RP-A and RAD51 at PML bodies
CBP	Protein acetyl-transferase, co-activator
Daxx	Involved in Fas-mediated apoptosis, transcriptional repression, and chromatin remodeling
Hipk2	Serine/threonine kinase; associates with p53 and CBP and is recruited to PML bodies
Mdm2	Regulates p53 protein levels
NBS1	Involved in DNA repair; complexes with Mre11 and Rad50 at PML bodies; complexes with Mre11 and TRF1 at PML bodies in ALT cells
p53	Tumour-suppressor, transcription factor
PML	Protein involved in several nuclear functions
Sp100	Transciptional repressor
SUMO-1	Small ubiquitin-like modifier (post-translational modification)
TRF1	Telomere-binding protein; colocalizes with PML bodies in ALT cells
TRF2	Telomere-binding protein; colocalizes with PML bodies in ALT cells

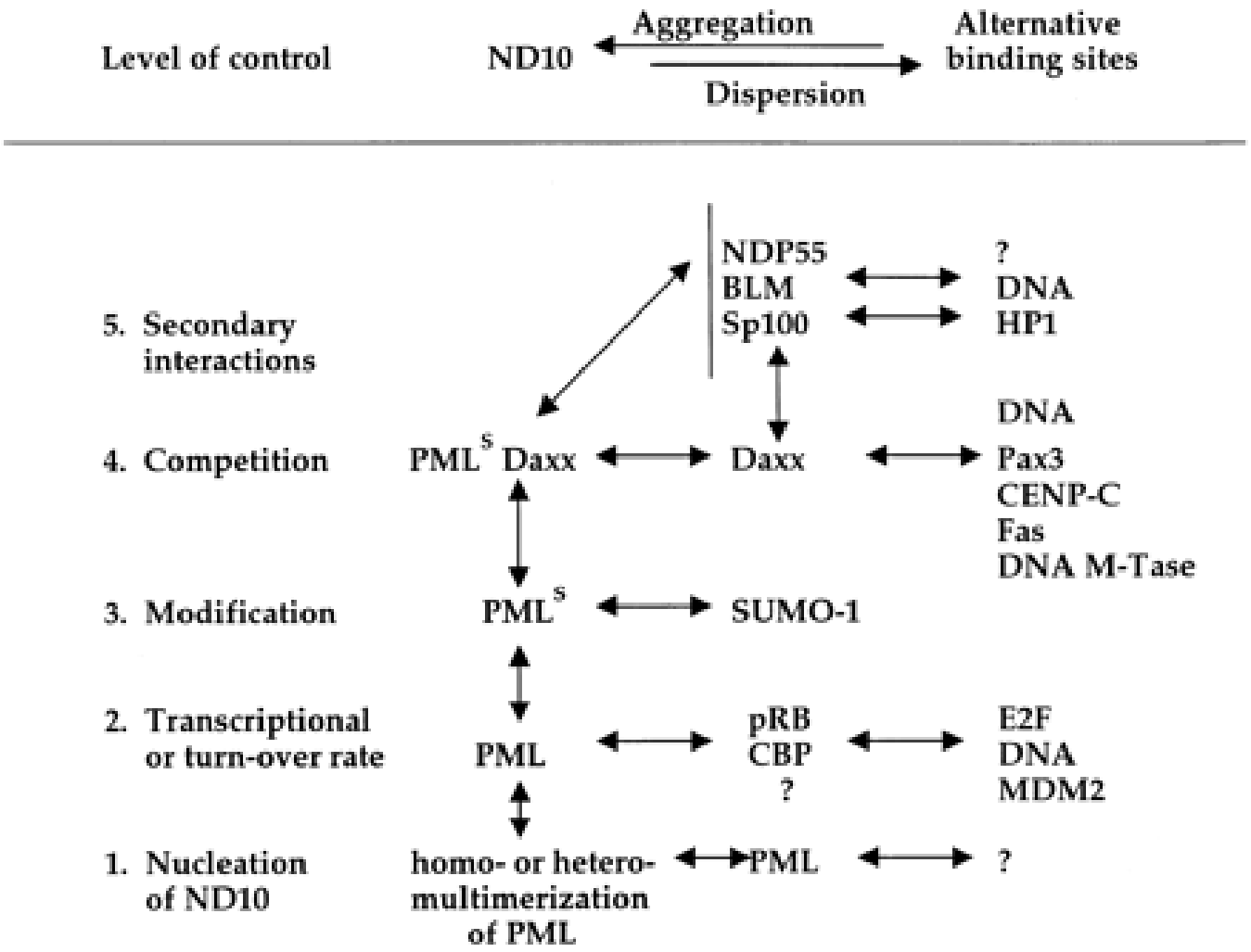
Ching et al. (2005) J. Cell Sci.

Formation

The appearance of ND10 after mitosis must result from a nucleation event possibly through homo- or heteromultimerization of PML. This event may take place at specific nuclear deposition sites... . Transcriptional activation, for instance by interferon, can upregulate PML expression ... nucleating additional aggregation sites. SUMO-1 modification-demodification of PML (third level) may lead to a reversible accumulation of Daxx to ND10 (fourth level), increasing or decreasing the availability of this protein for alternative binding partners (DNA, CENP-C, Fas, Pax3, DNA methyltransferase), and thus regulate corresponding functions. The complexity and plasticity of such a supramolecular regulatory mechanism are evident and envisioned structurally as a network of interacting proteins with PML at its core.

Ishov et. al. (1999), J. Cell. Biol.

Formation



Function

PML bodies are implicated in many cellular processes

- tumour suppression
- apoptosis,
- DNA replication and repair,
- proteolysis and response to viral infection
- nuclear protein storage,
- gene regulation and transcription,

PML protein is implicated in acute pro-myelocytic leukaemia (APL)

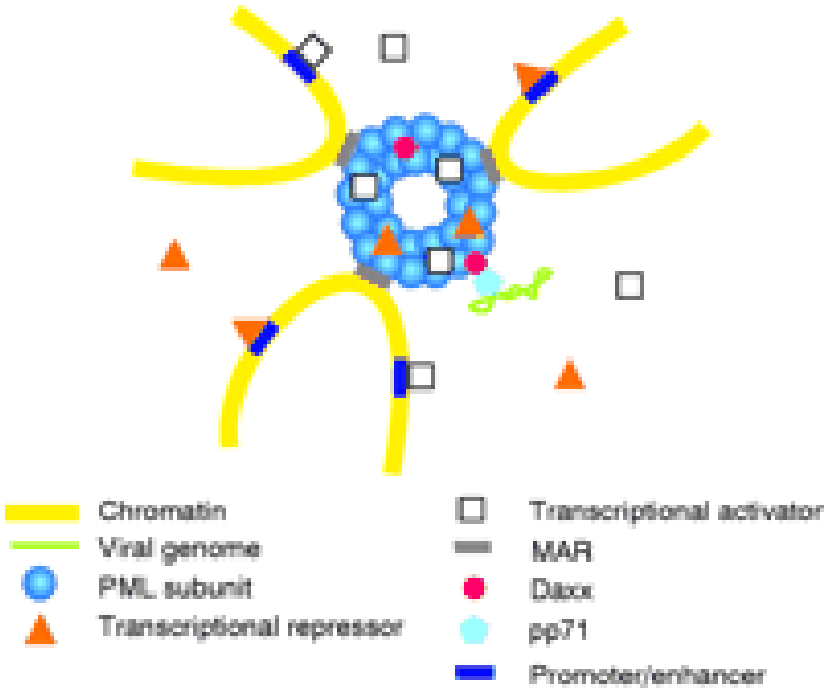
Putative Functional Model

PML bodies consist of smaller PML-containing subunits that interact with high-order chromatin fibers possibly through matrix attachment regions (MARs). PML bodies function to control the concentration of transcriptional activators and repressors within the local chromatin environment. Upon stimulation of cellular pathways, these factors are released from PML bodies and bind enhancer/promoter elements in the surrounding chromatin. Upon infection with double-stranded DNA viruses such as HCMV, viral genomes localize to the surface of PML bodies. This targeting is mediated through cellular factors such as Daxx, which are normally associated with PML bodies, and viral tegument proteins, such as pp71.

Ching et. al. (2005), J. Cell. Sci.

Schematic

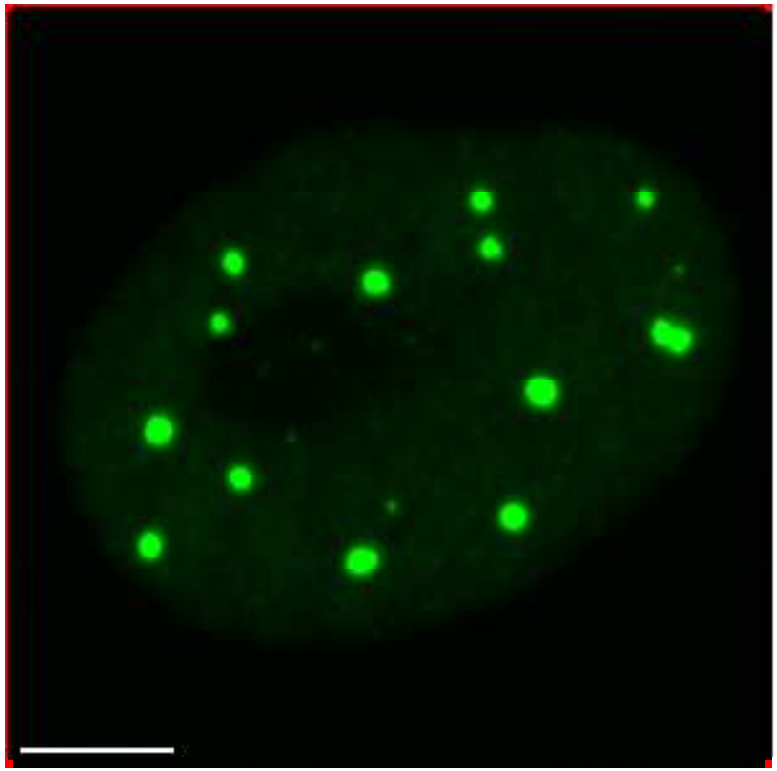
Schematic representation of PML bodies in the nucleoplasm.



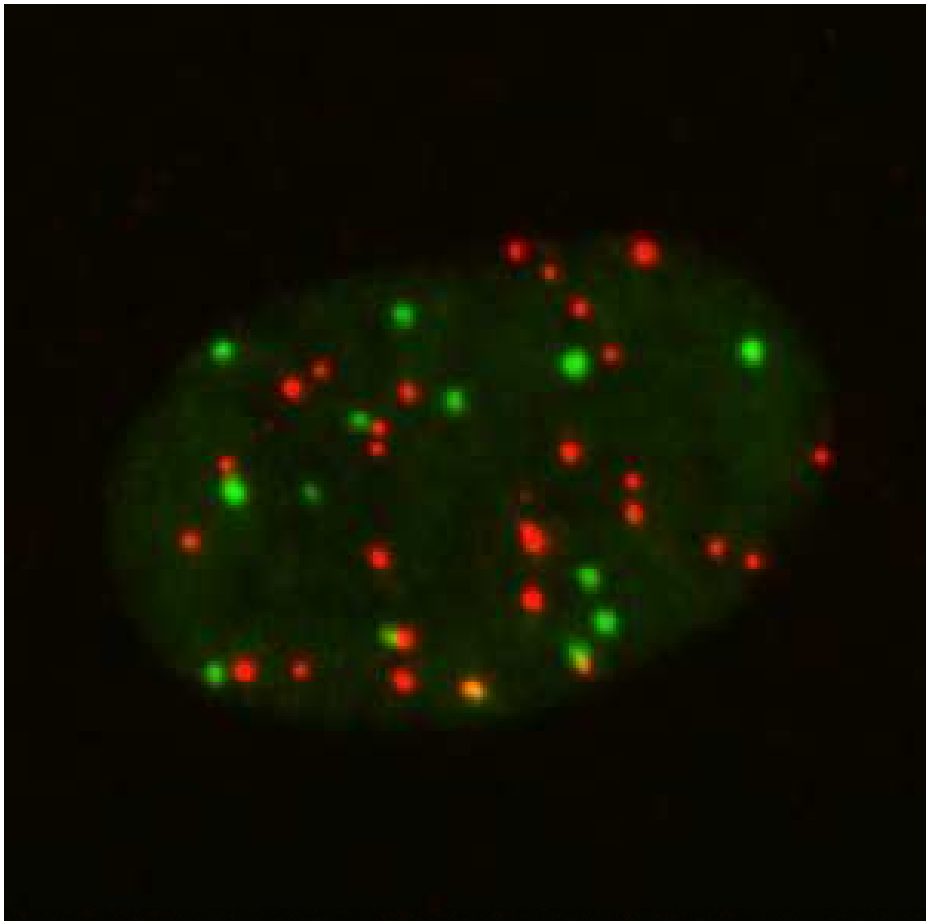
Imaging (Batty, 2006)

Nuclear compartments imaged using immunofluorescence staining and confocal microscopy

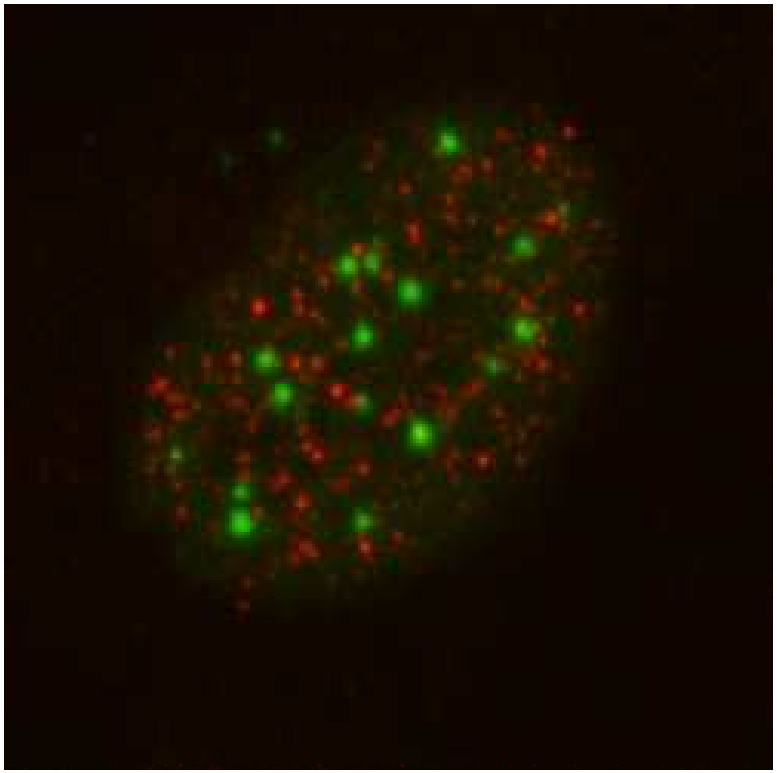
PML bodies



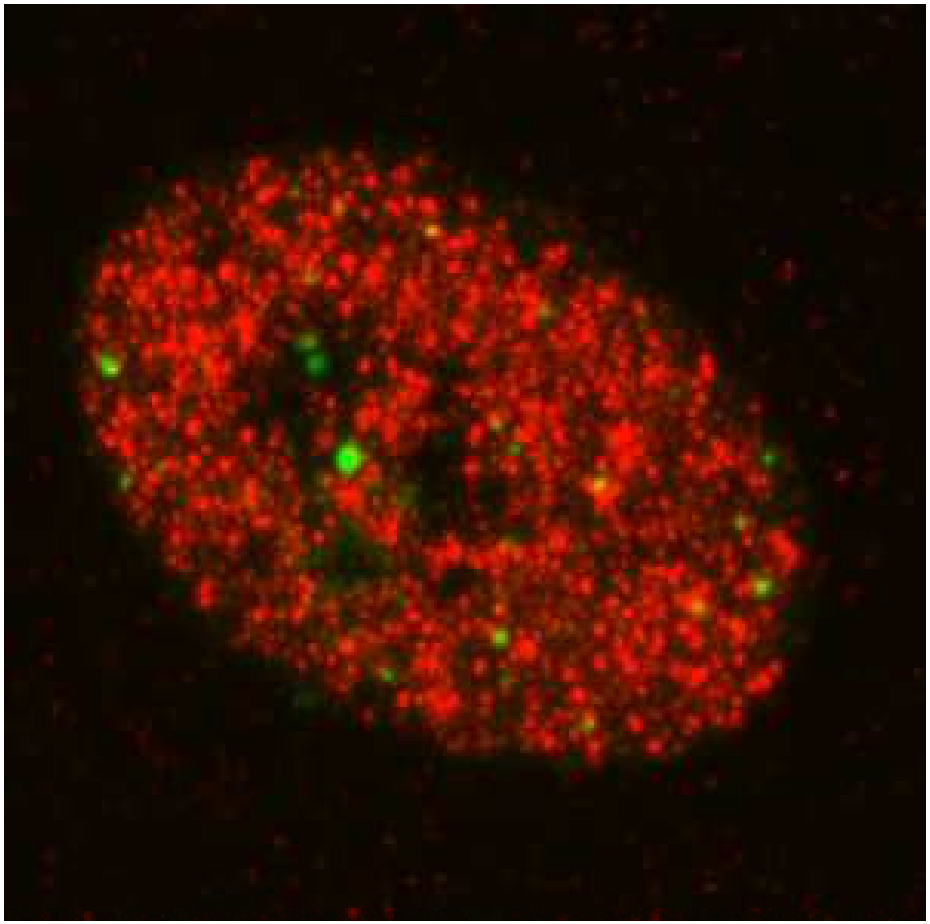
Centromeres



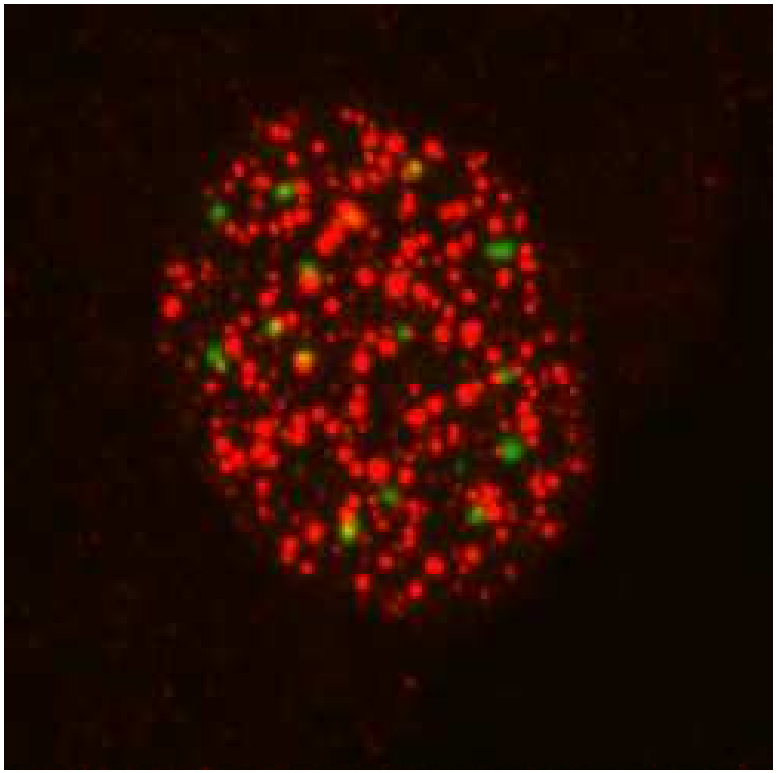
Nascent RNA



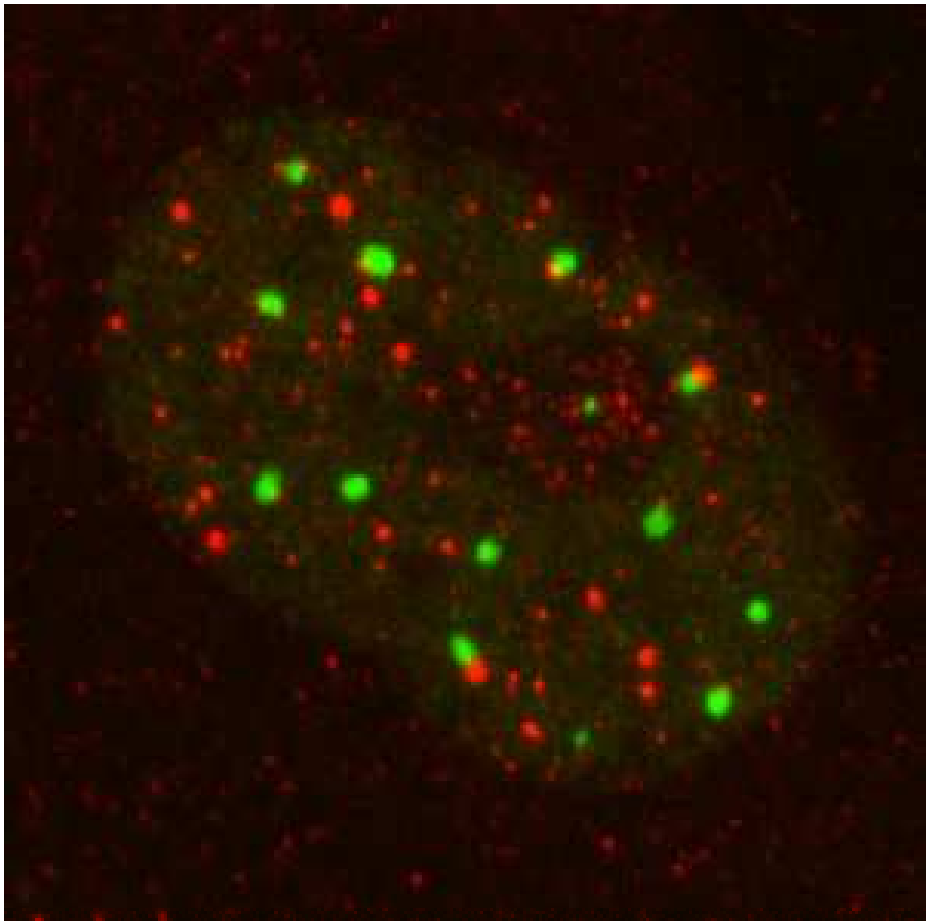
RNA polymerase II



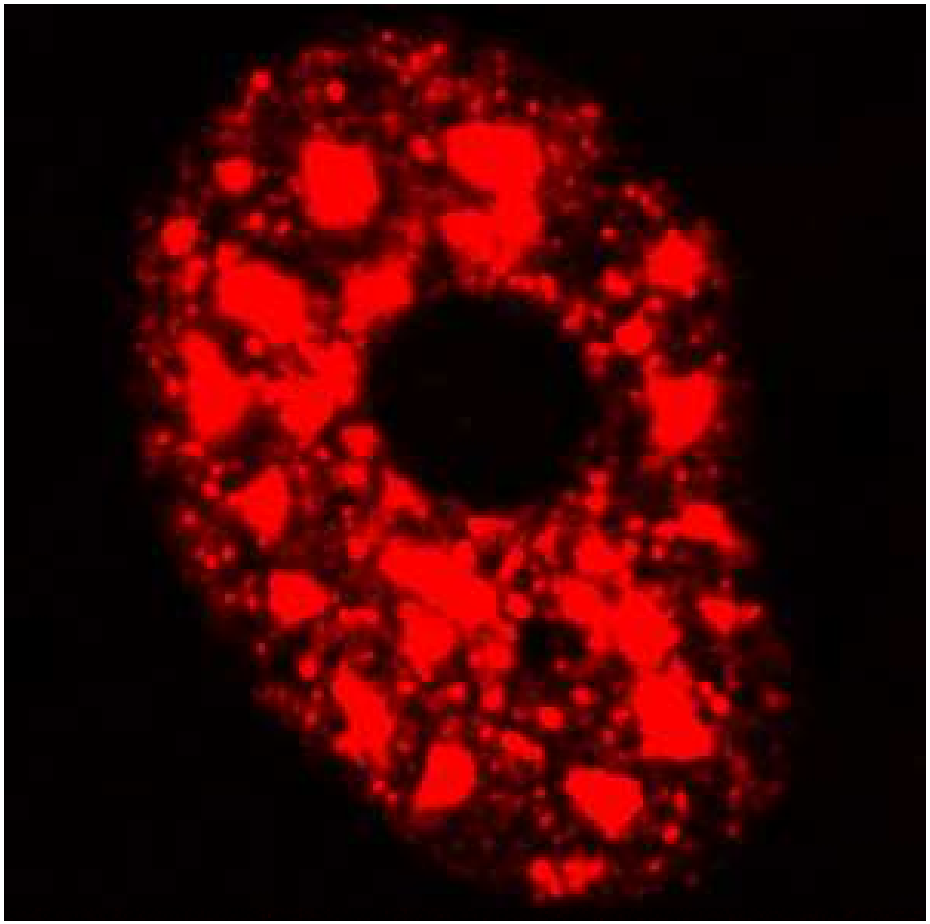
Acetylated histone



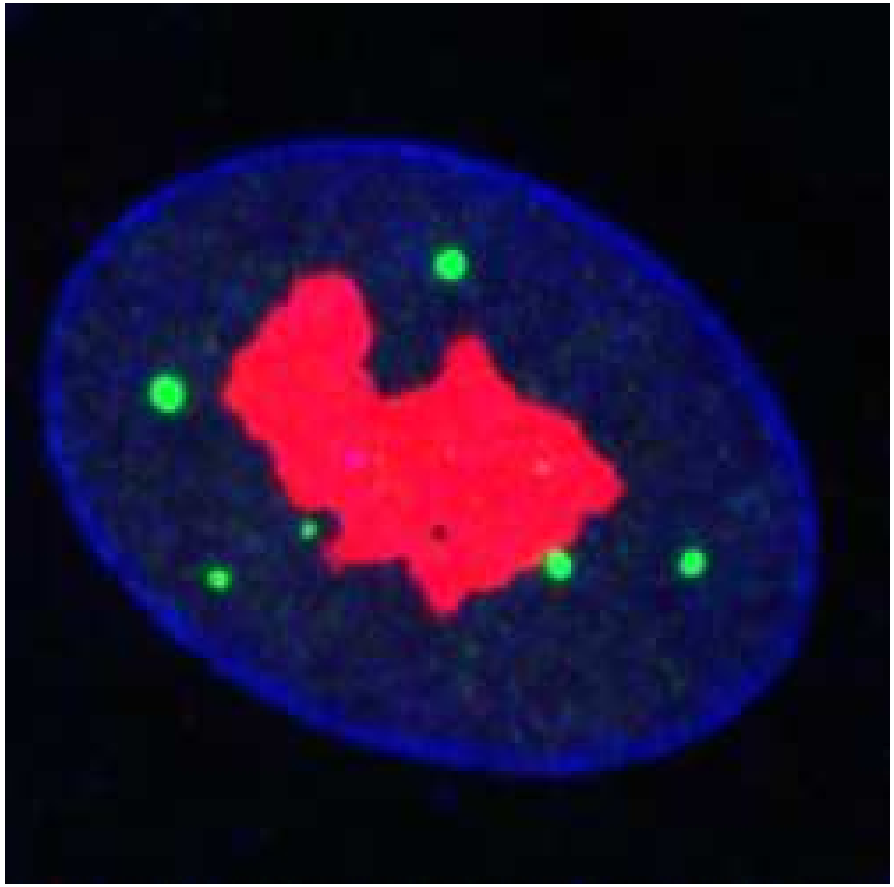
19S core proteasome



Splicing speckles



Nucleolus (red), lamin B of nuclear lamina (blue)



PML Bodies are Positionally Stable

PML bodies are positionally stable over long periods during interphase. Following stress, they are conserved in size and position after a cycle of disruption and re-formation. We conclude that their position and size is non-random and might be dictated by chromatin, which might restrict their mobility.

Ching et al. (2005), J. Cell Sci.

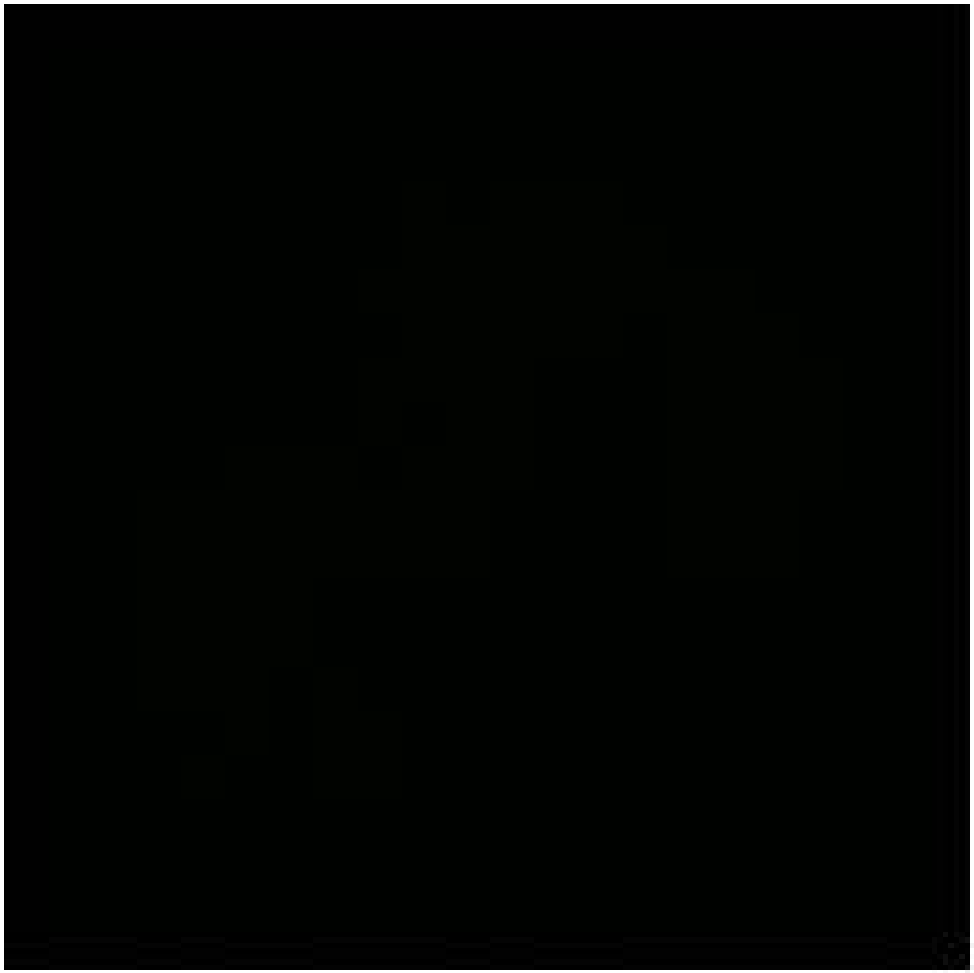
PML Bodies are Positionally Stable

Bazett-Jones Lab, U Toronto
http://www.sickkids.ca/bazett-joneslab/

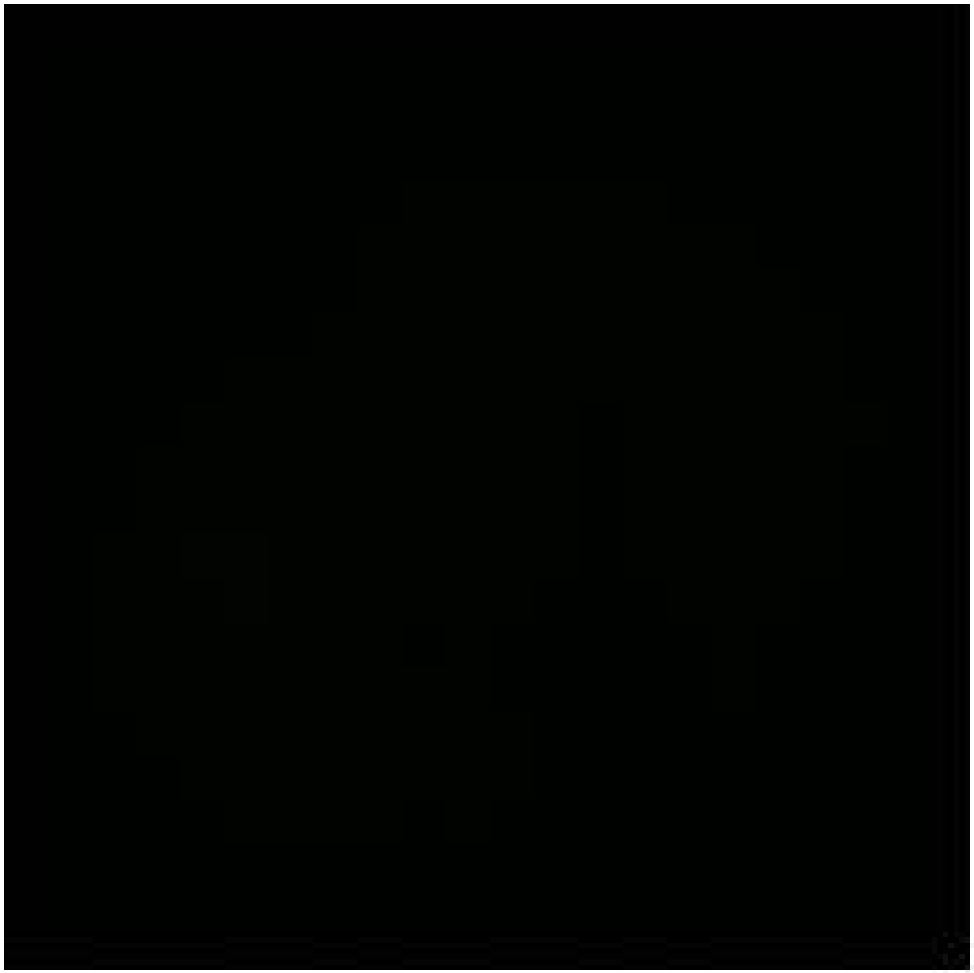
Evidence of Dynamic Colocalization

Daniels et al. (2004) Nat. Struc. & Mol. Biol.

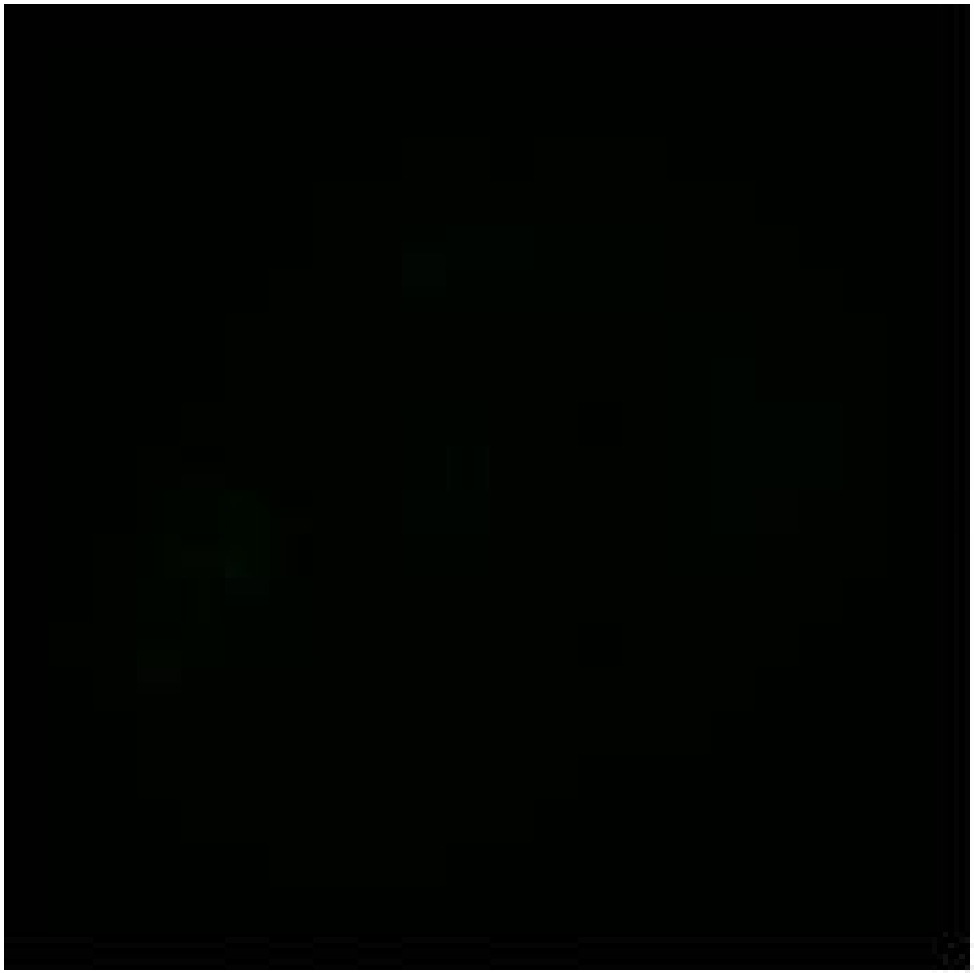
Slices: PML



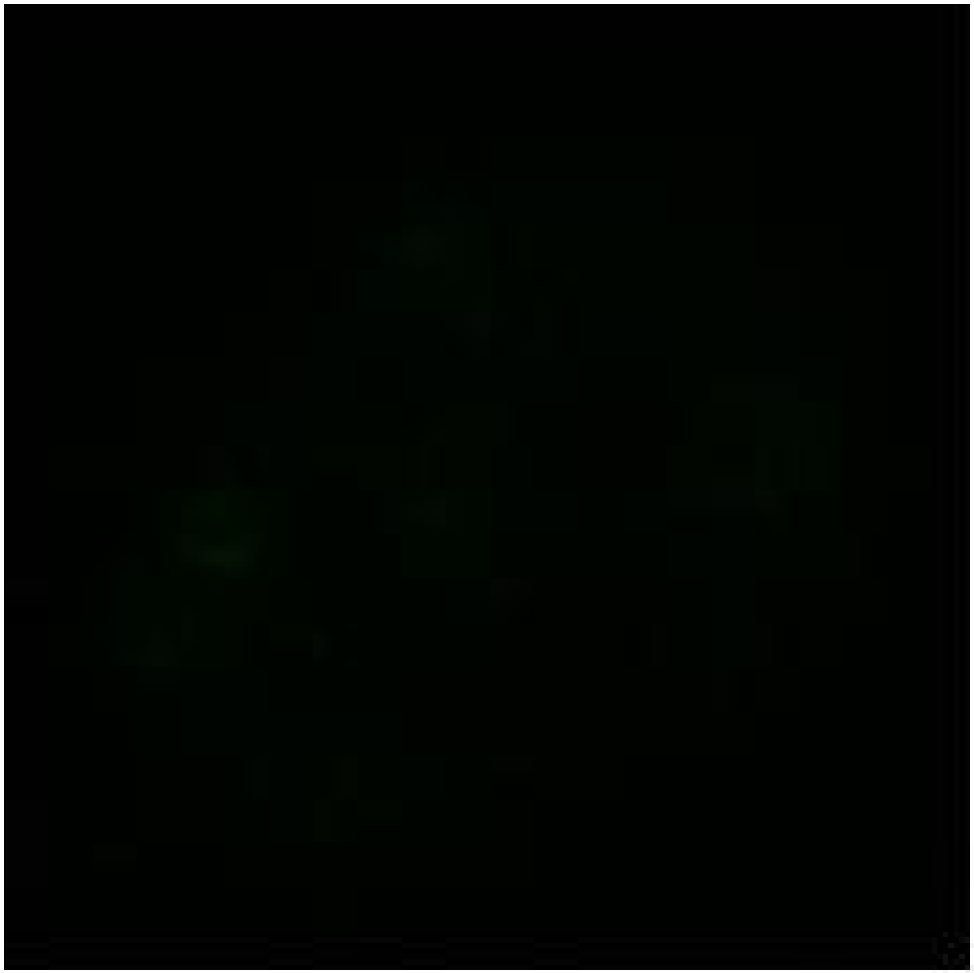
Slices: PML



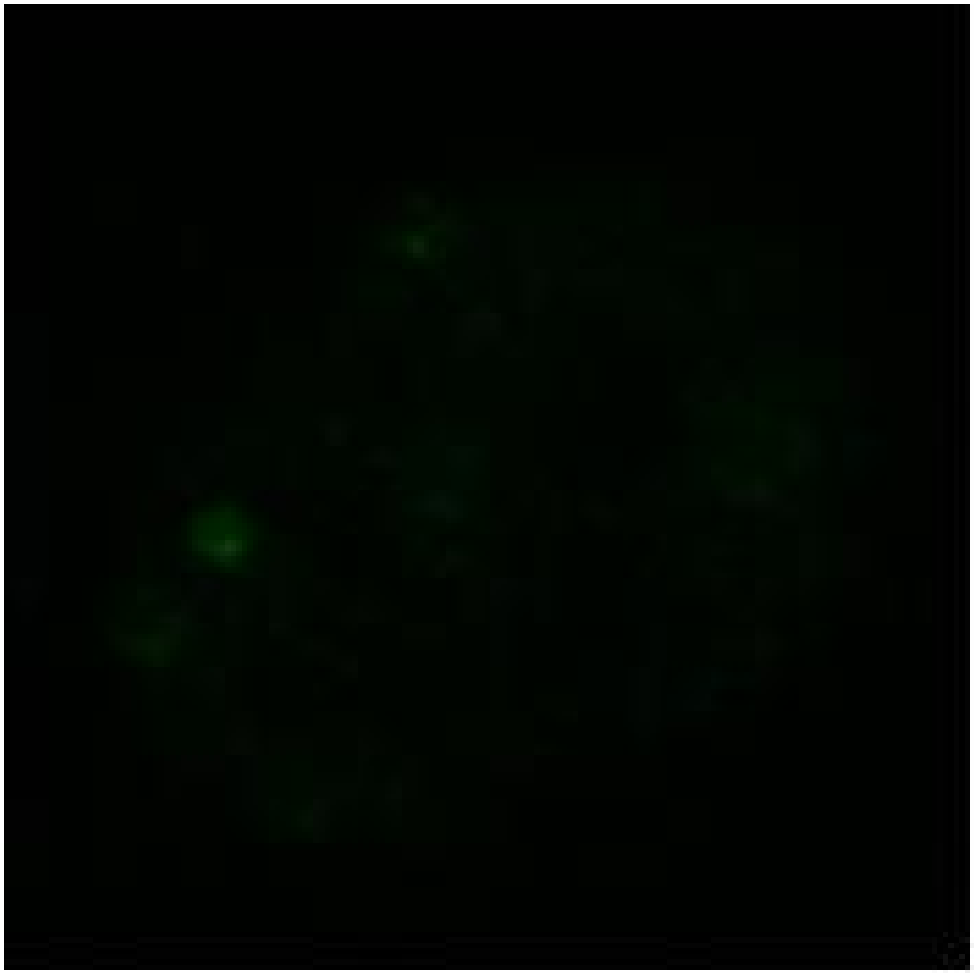
Slices: PML



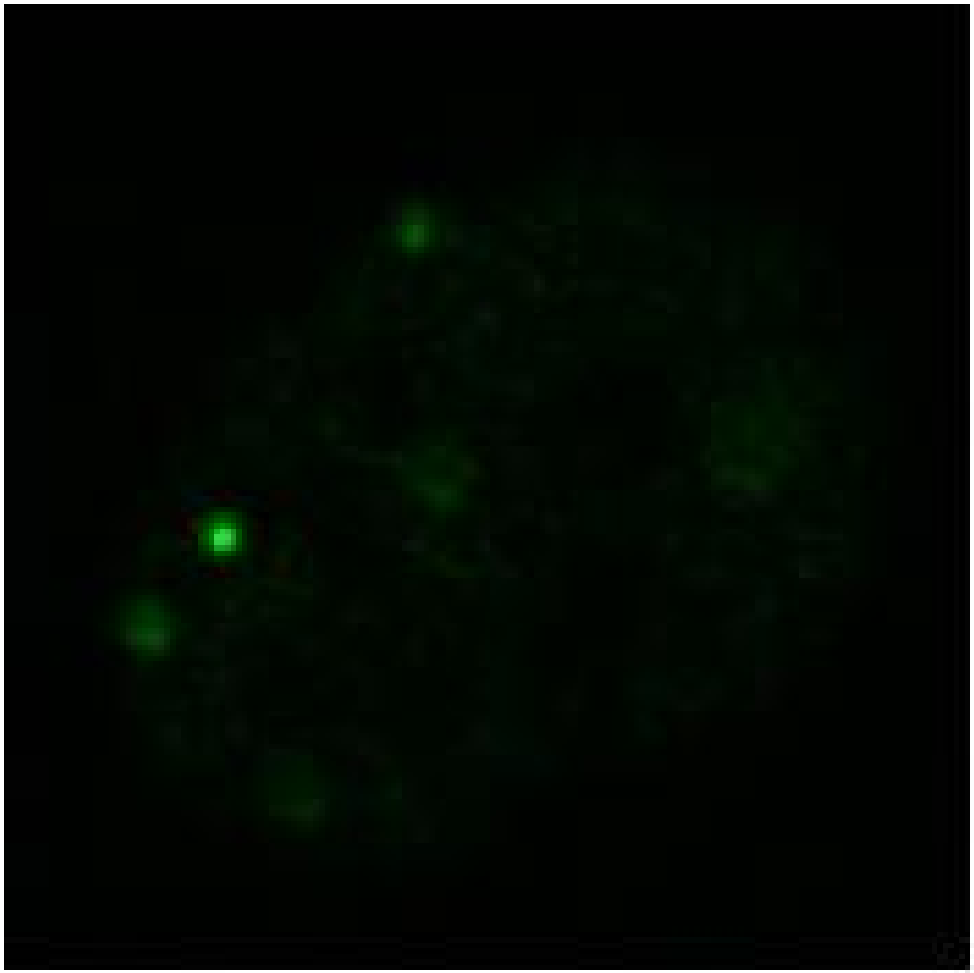
Slices: PML



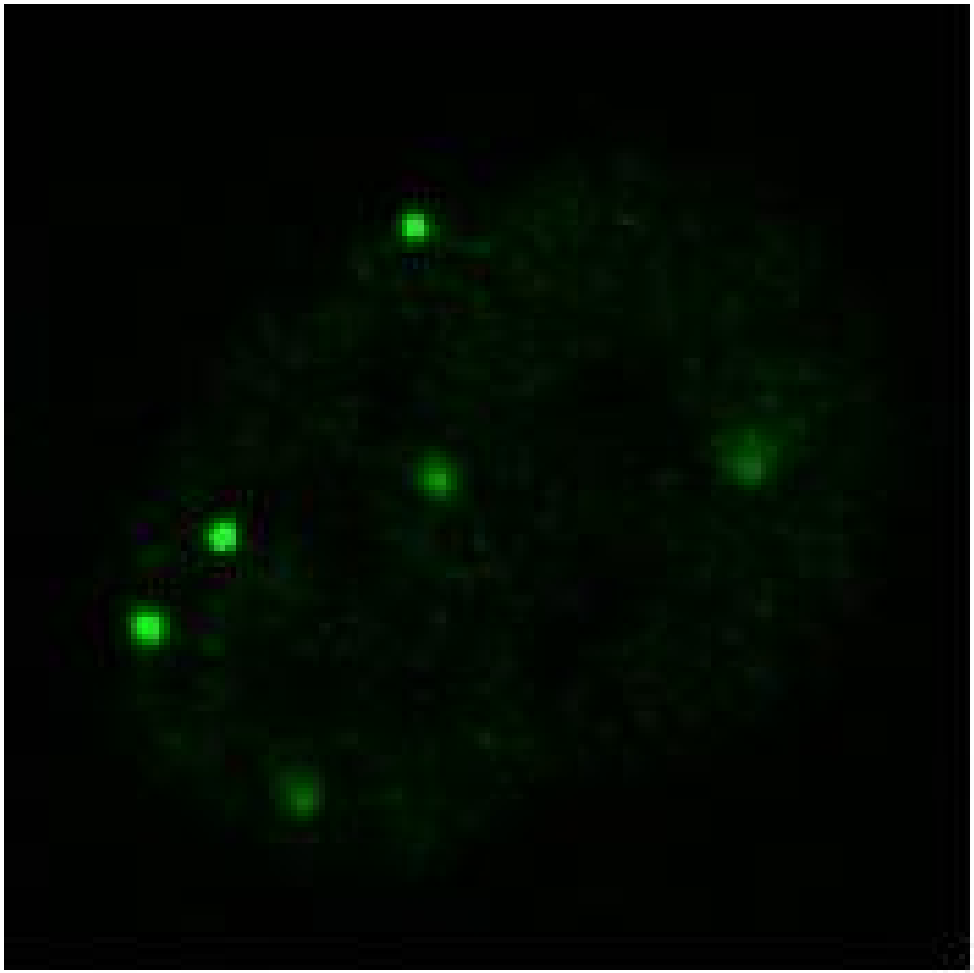
Slices: PML



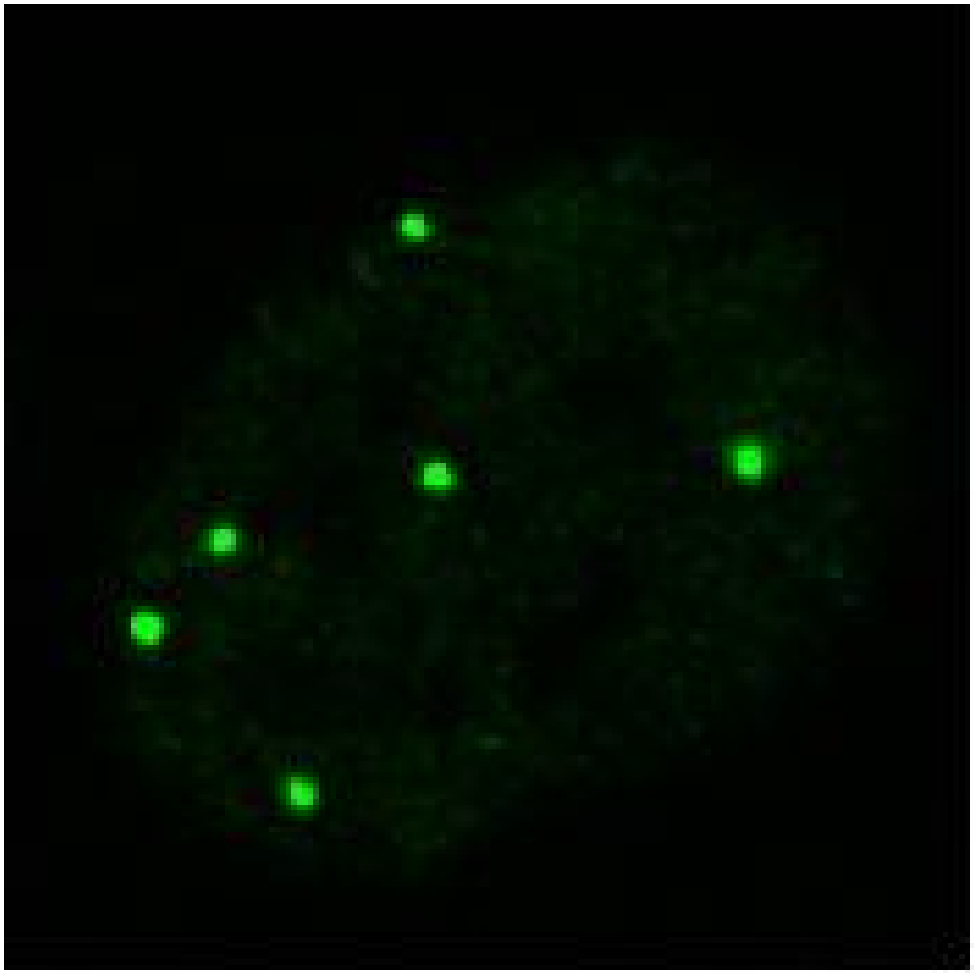
Slices: PML



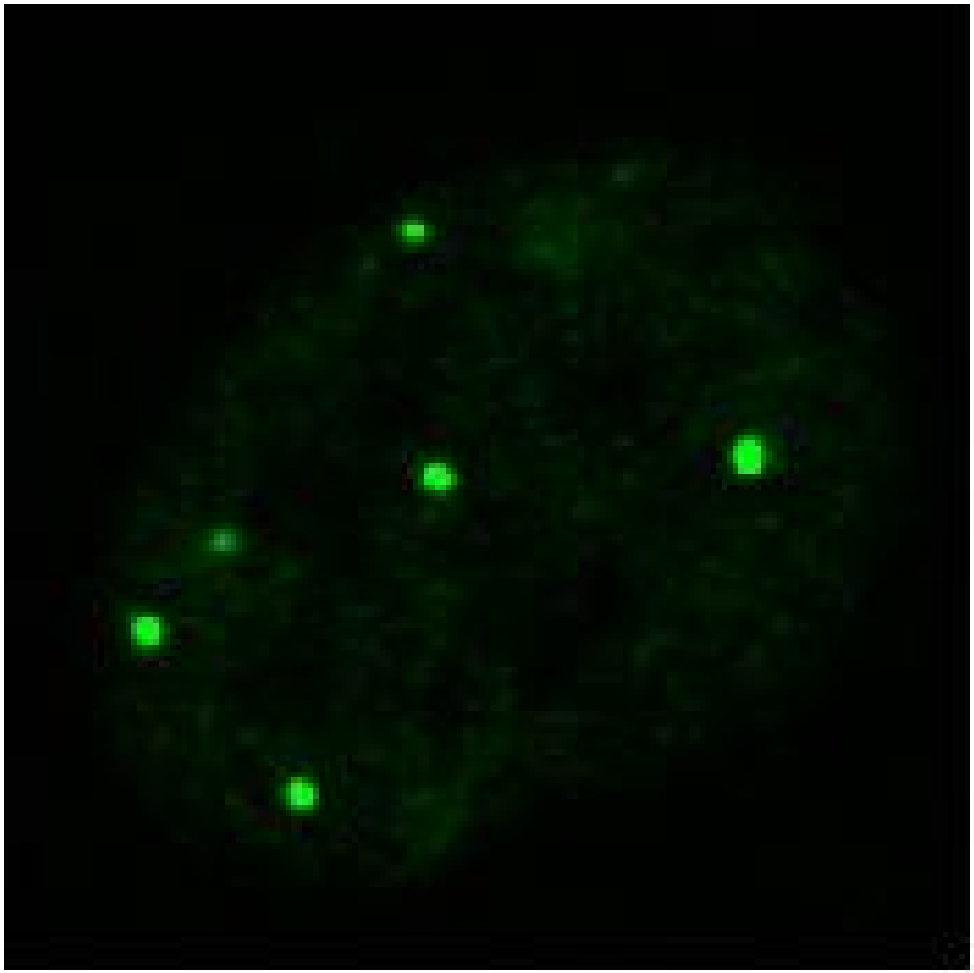
Slices: PML



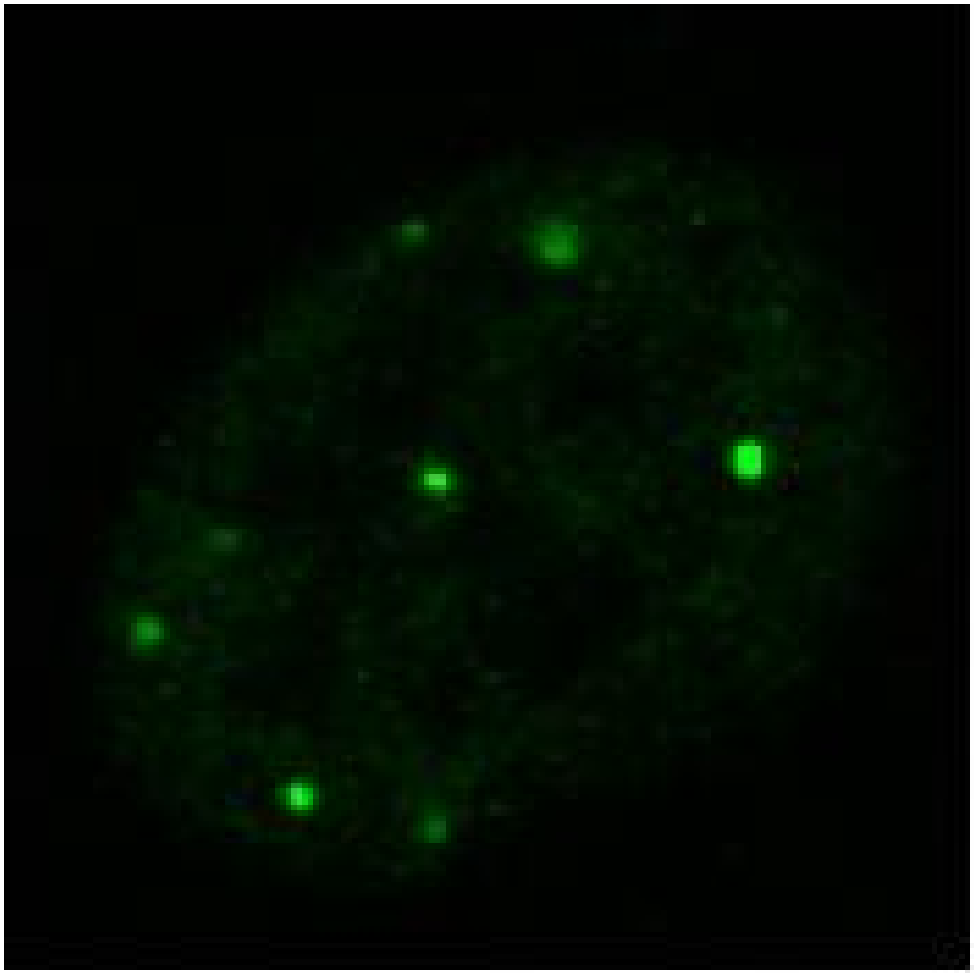
Slices: PML



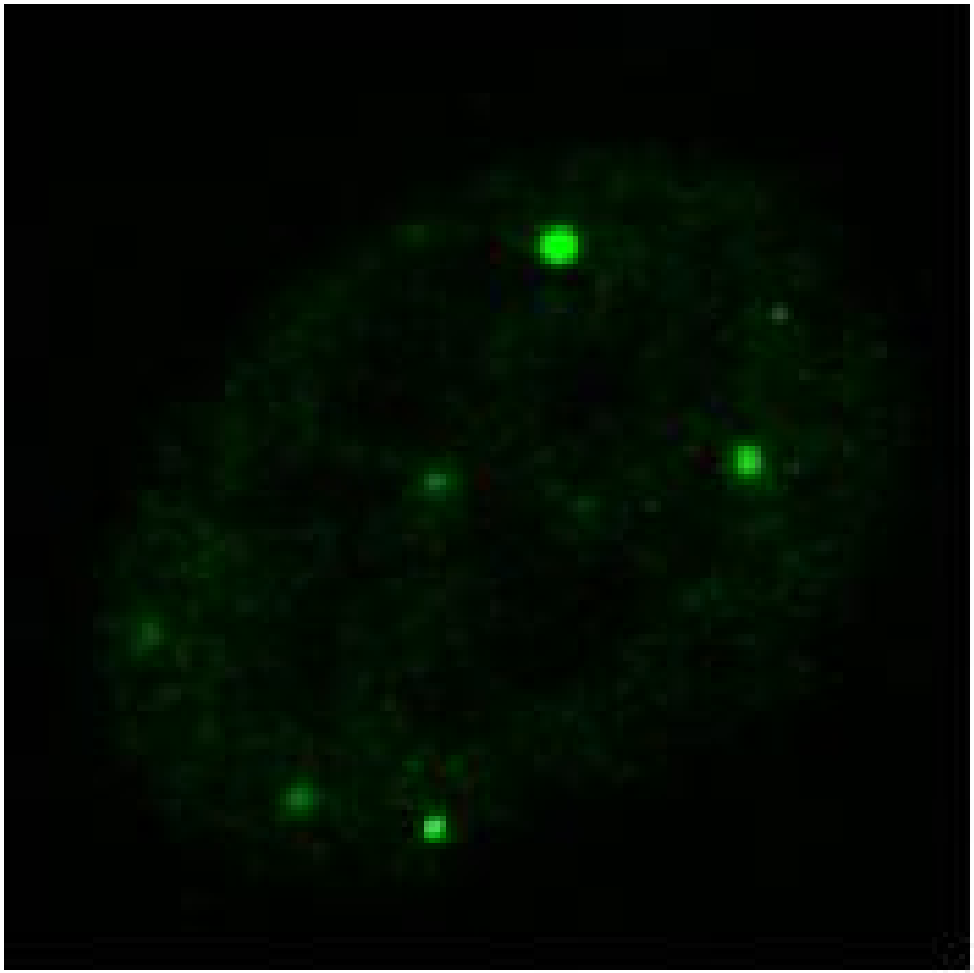
Slices: PML



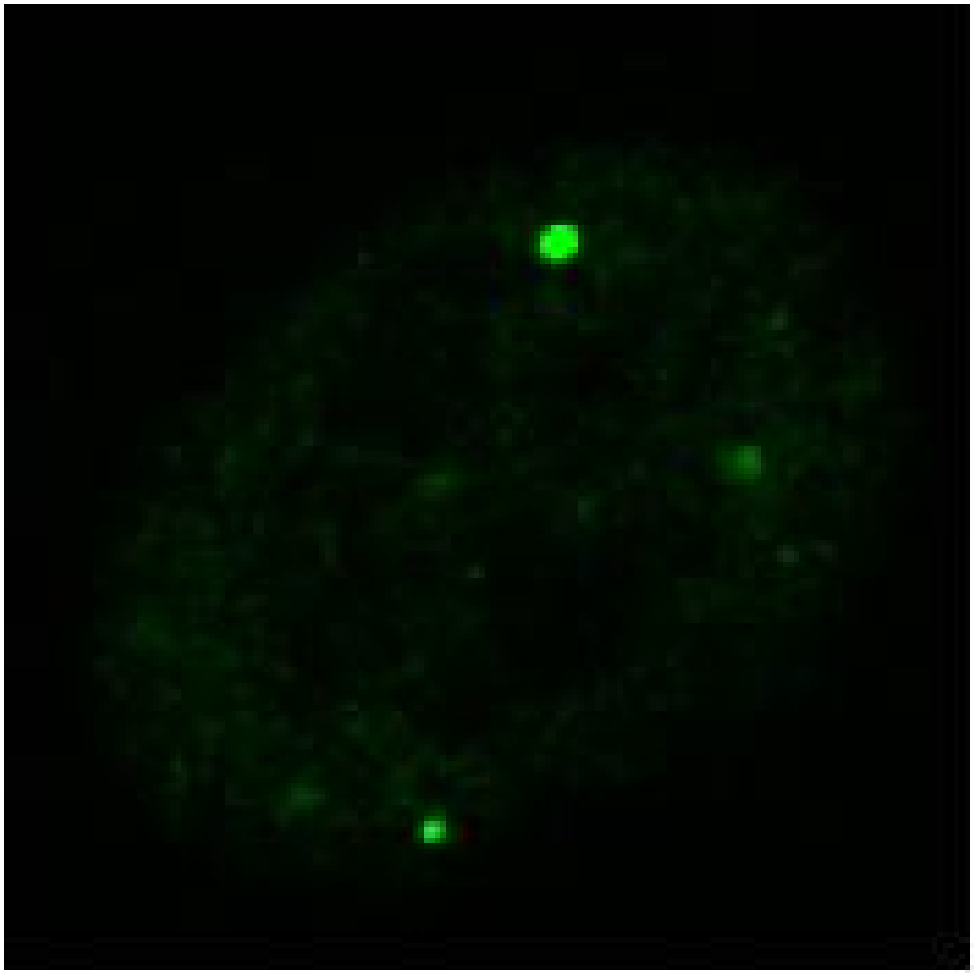
Slices: PML



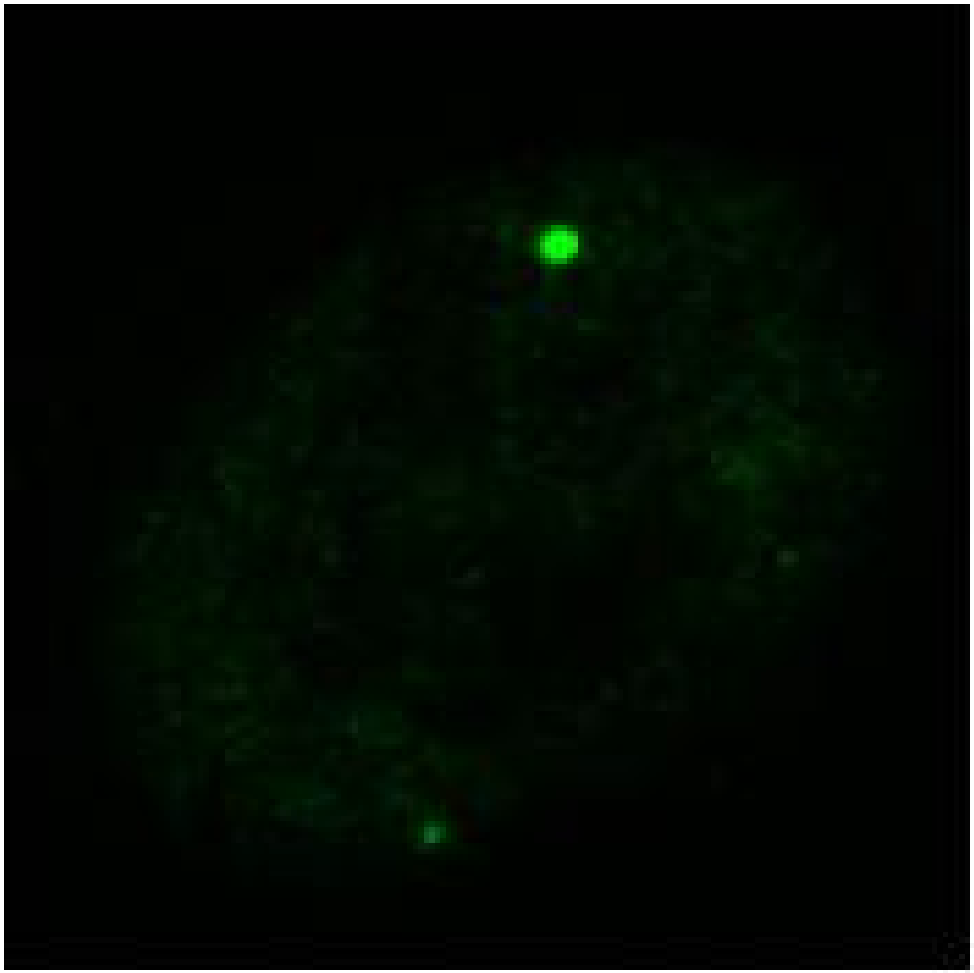
Slices: PML



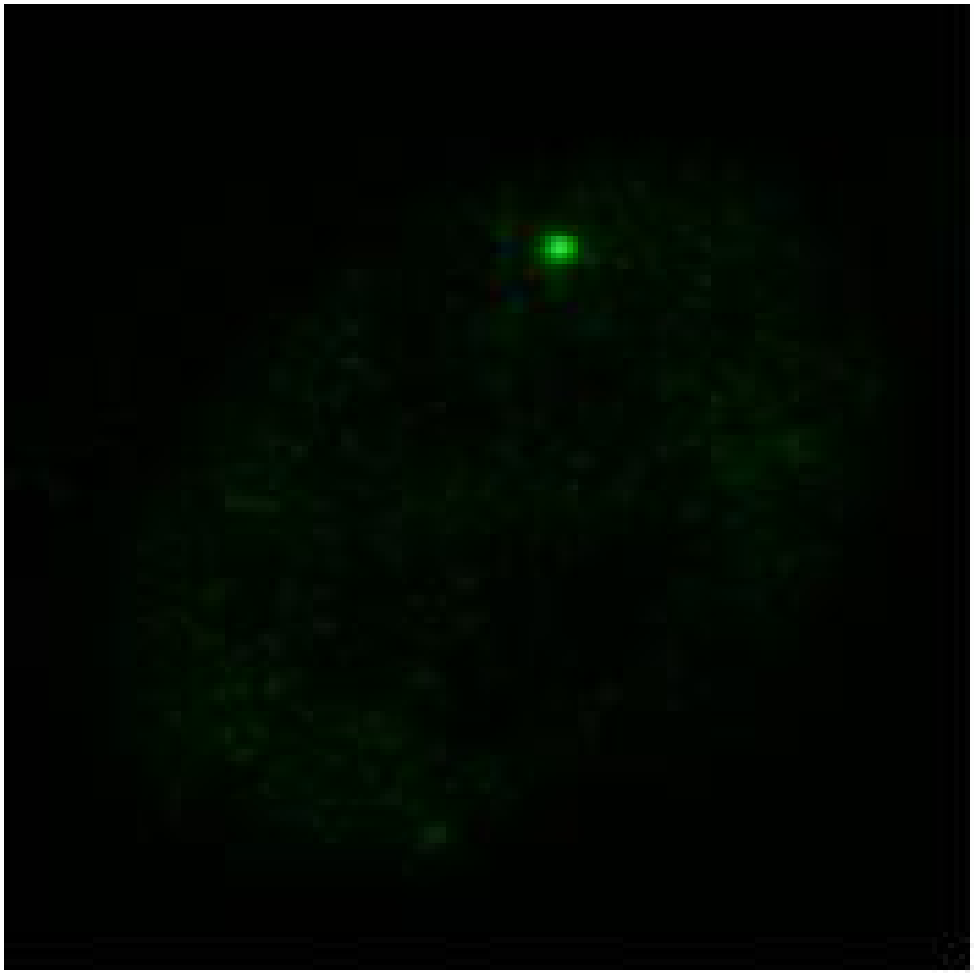
Slices: PML



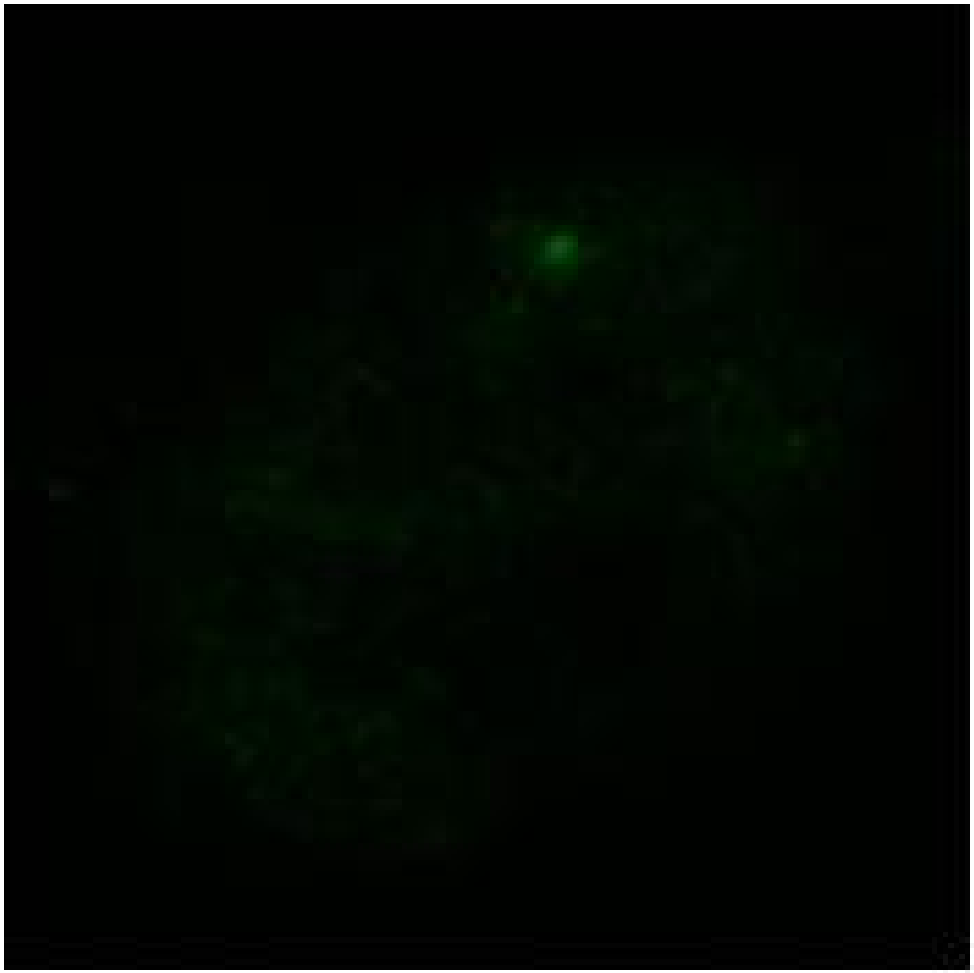
Slices: PML



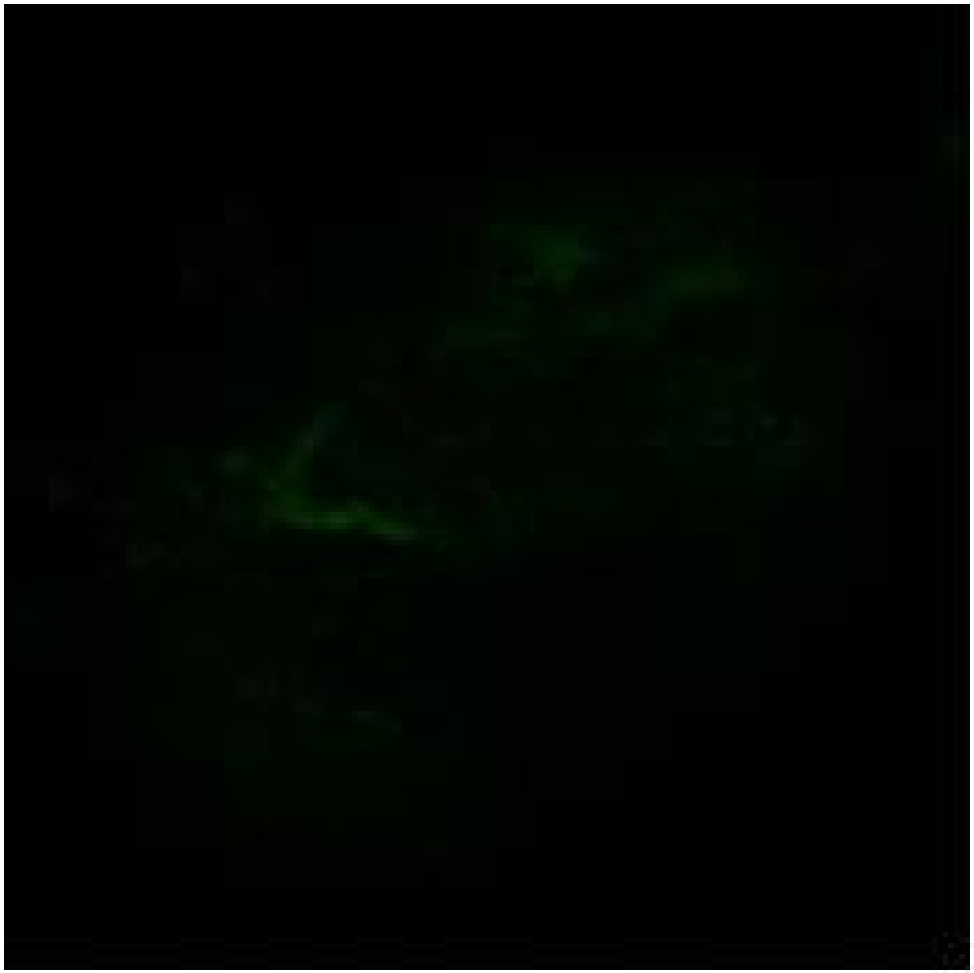
Slices: PML



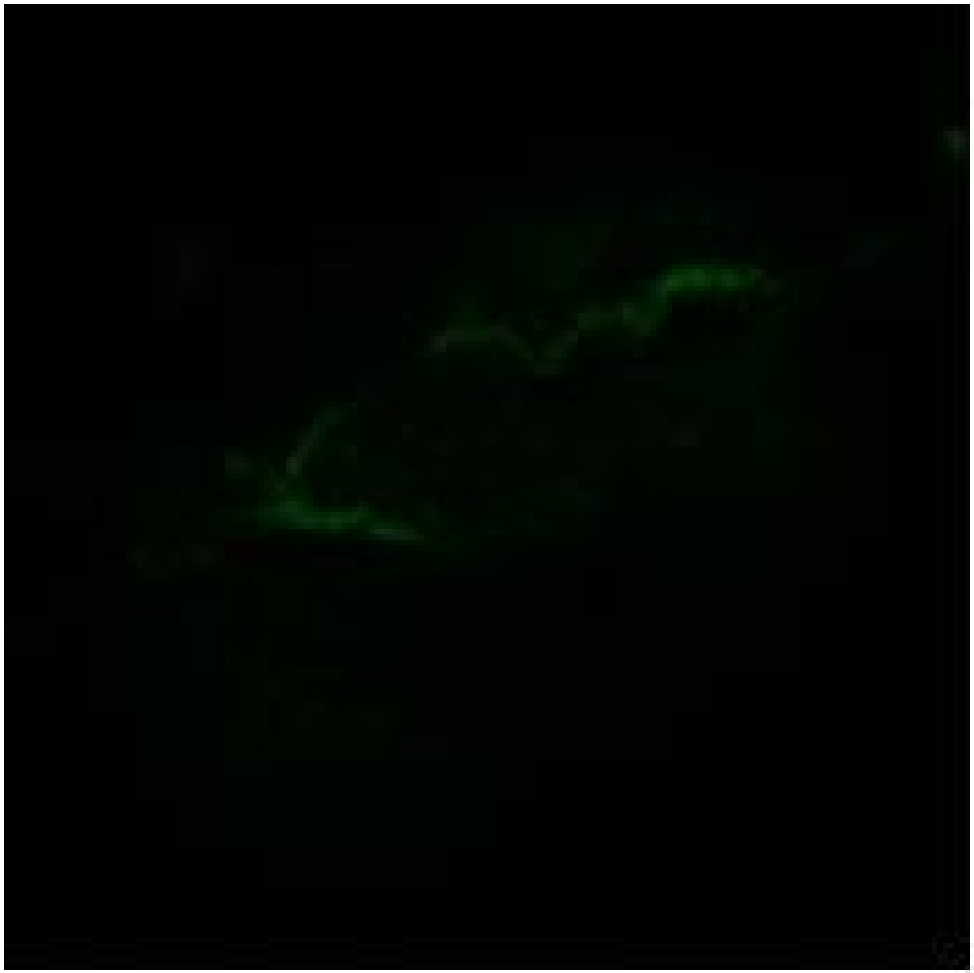
Slices: PML



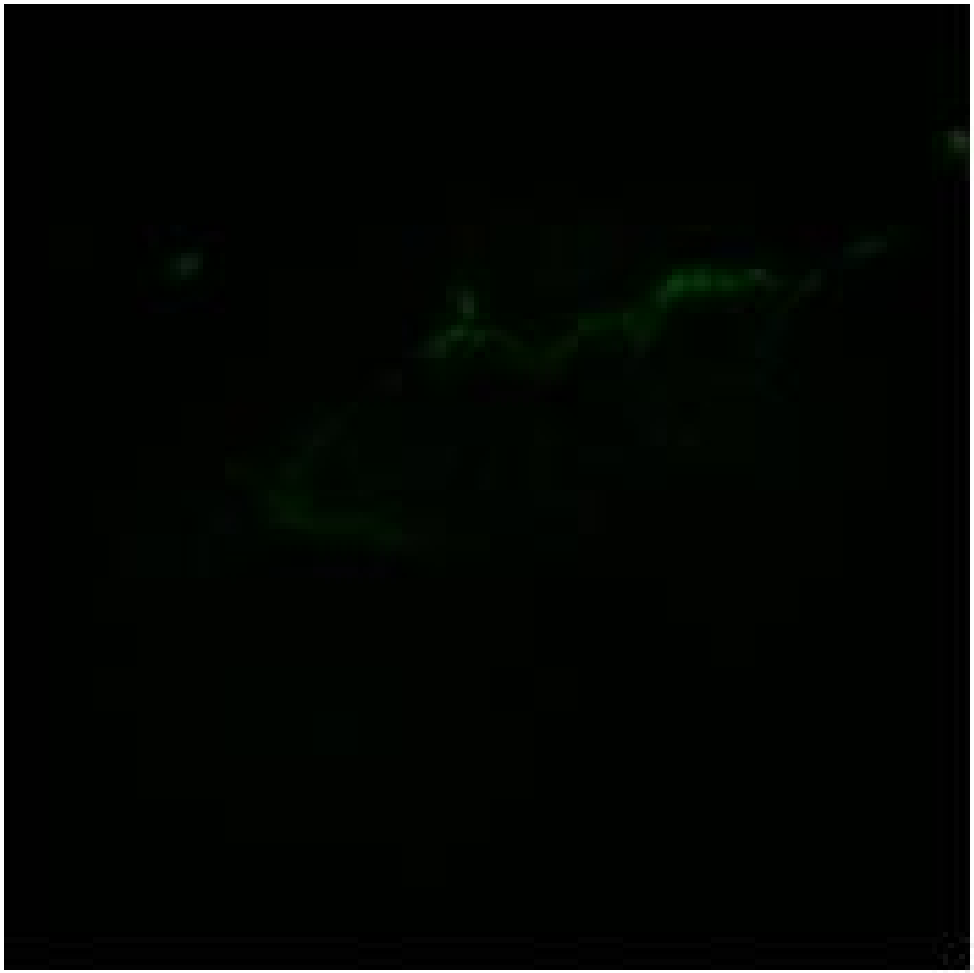
Slices: PML



Slices: PML

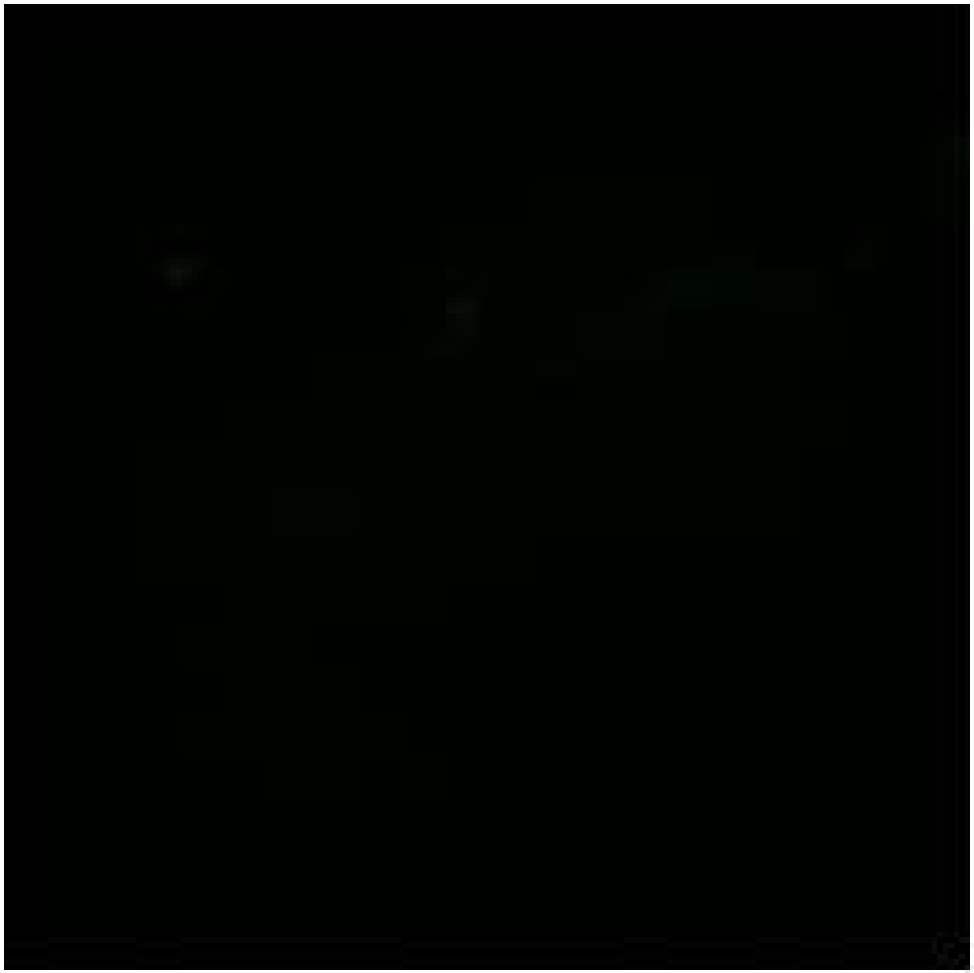


Slices: PML

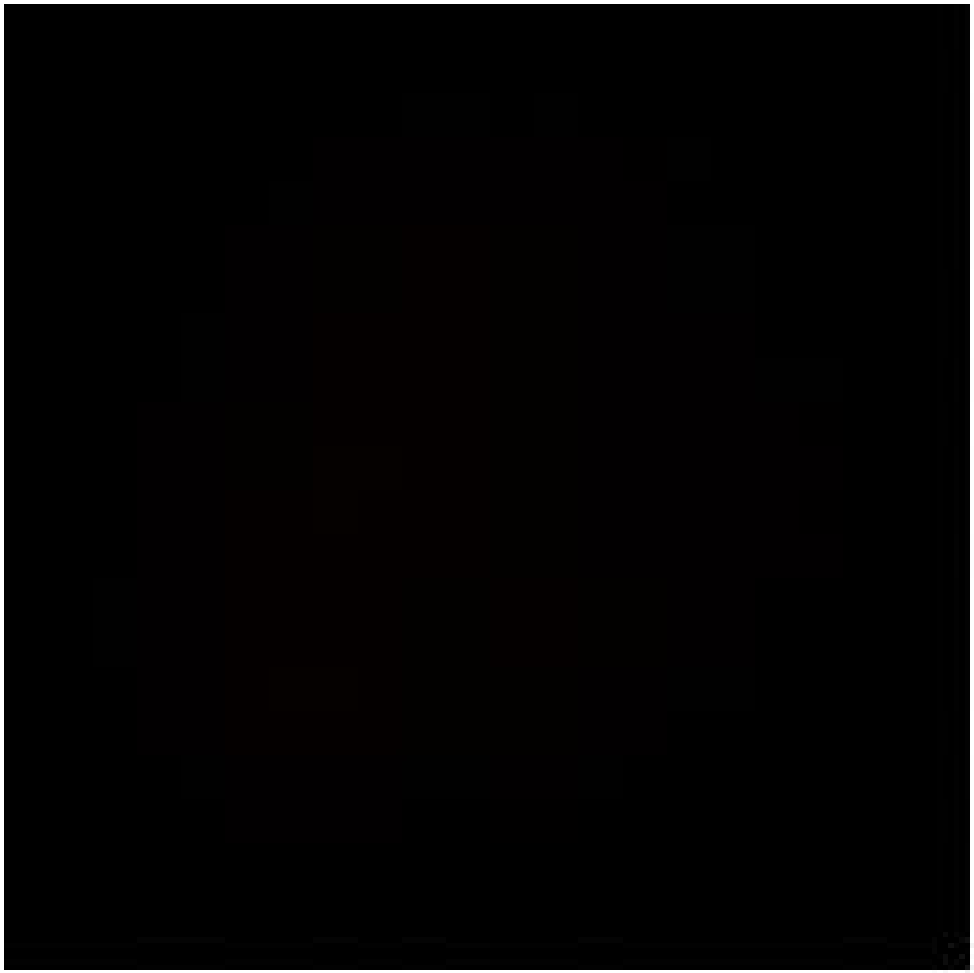


Slices: PML

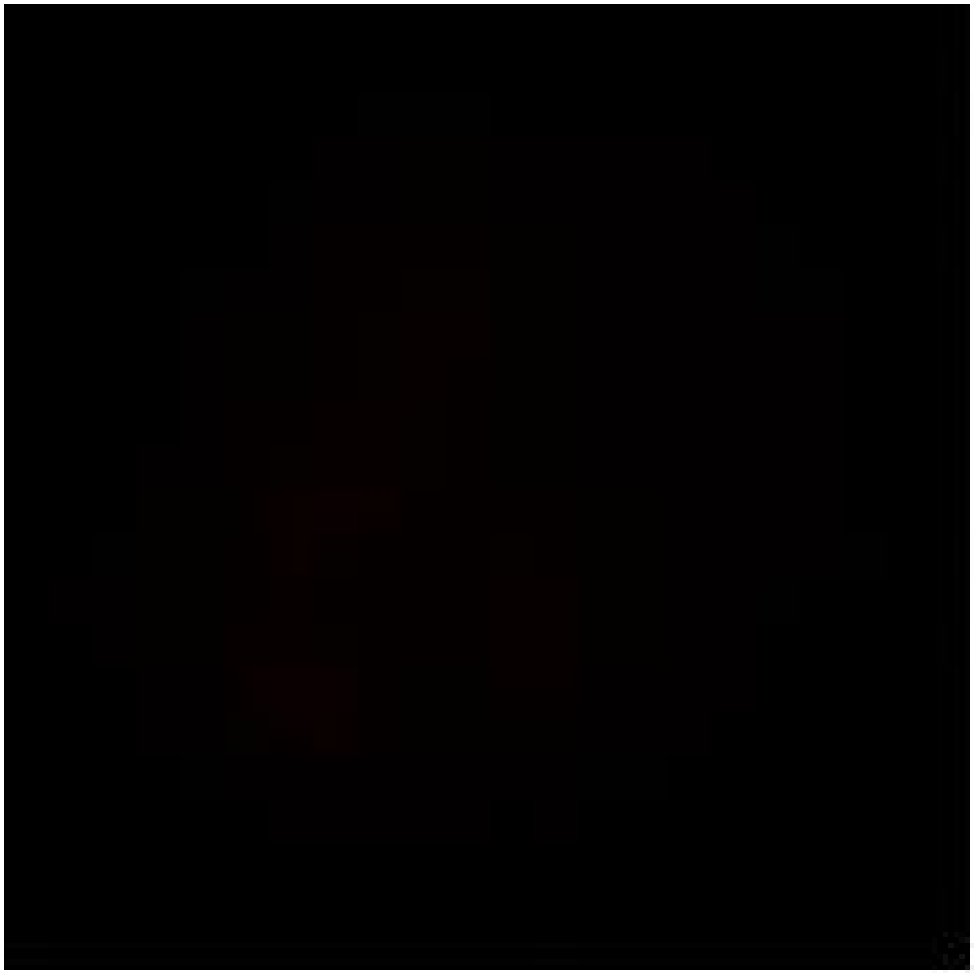
Slices: PML



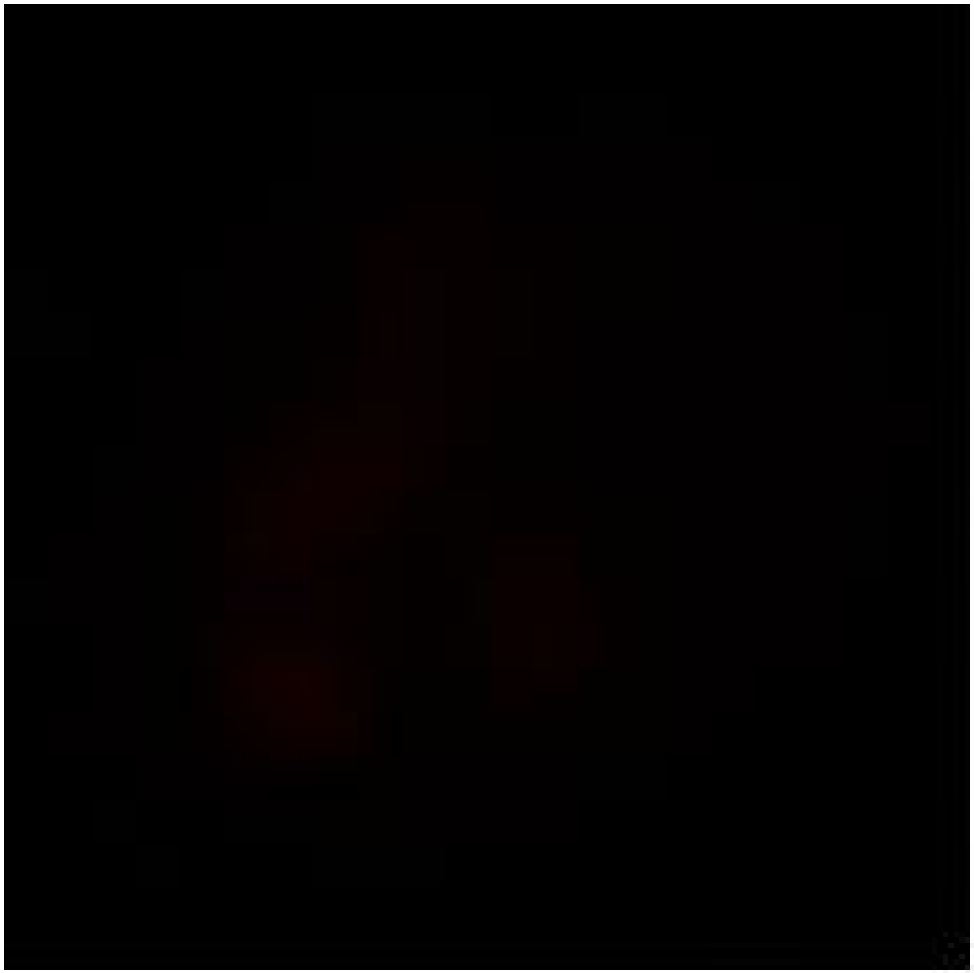
Slices: Nucleoli



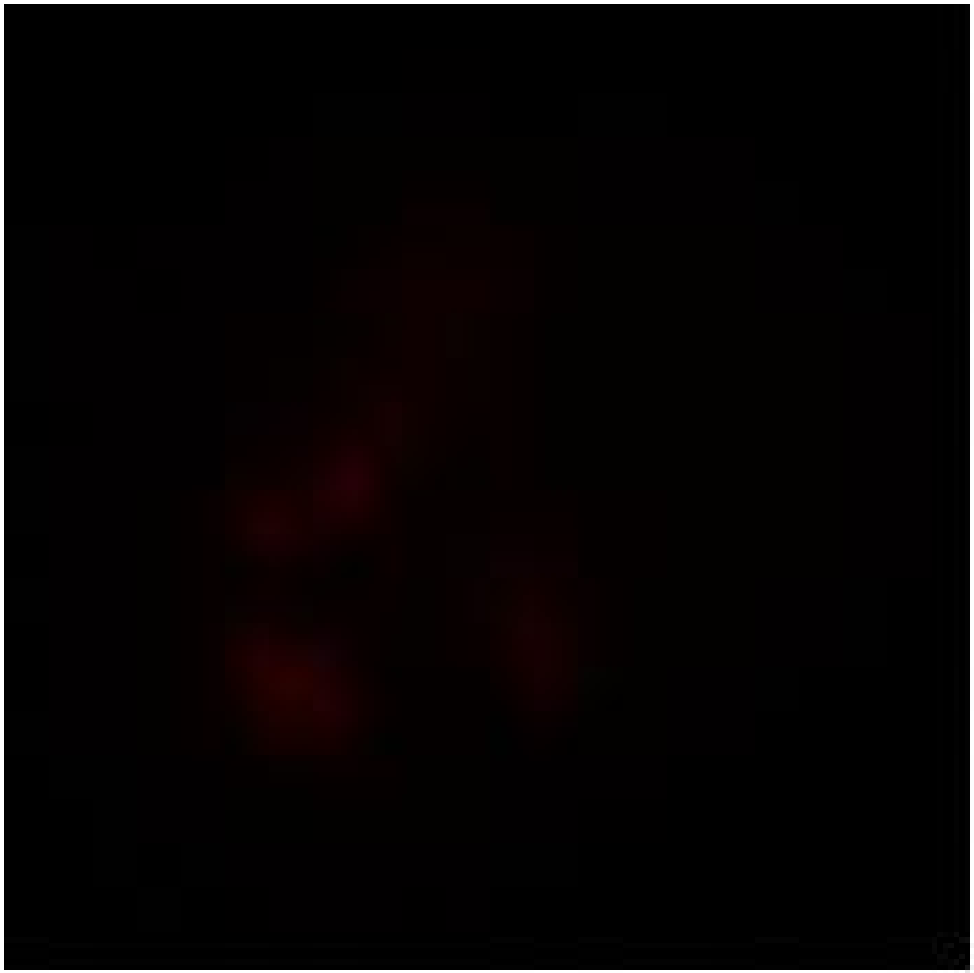
Slices: Nucleoli



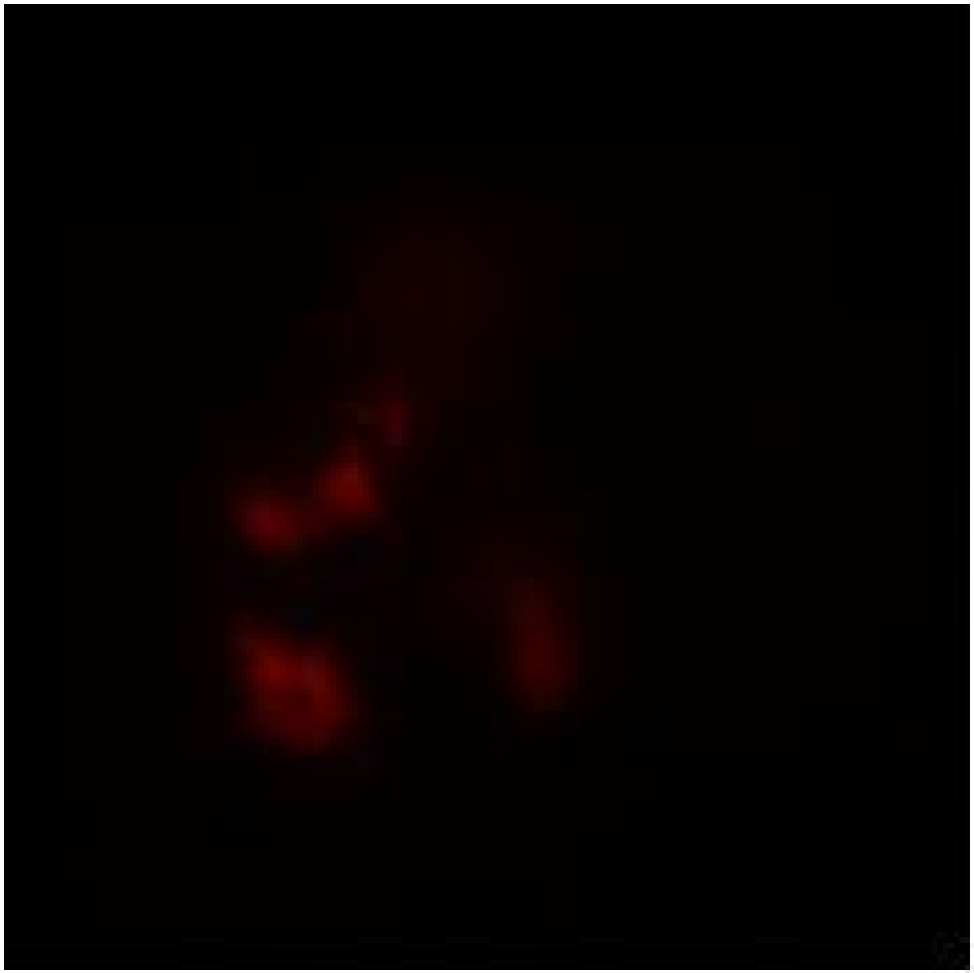
Slices: Nucleoli



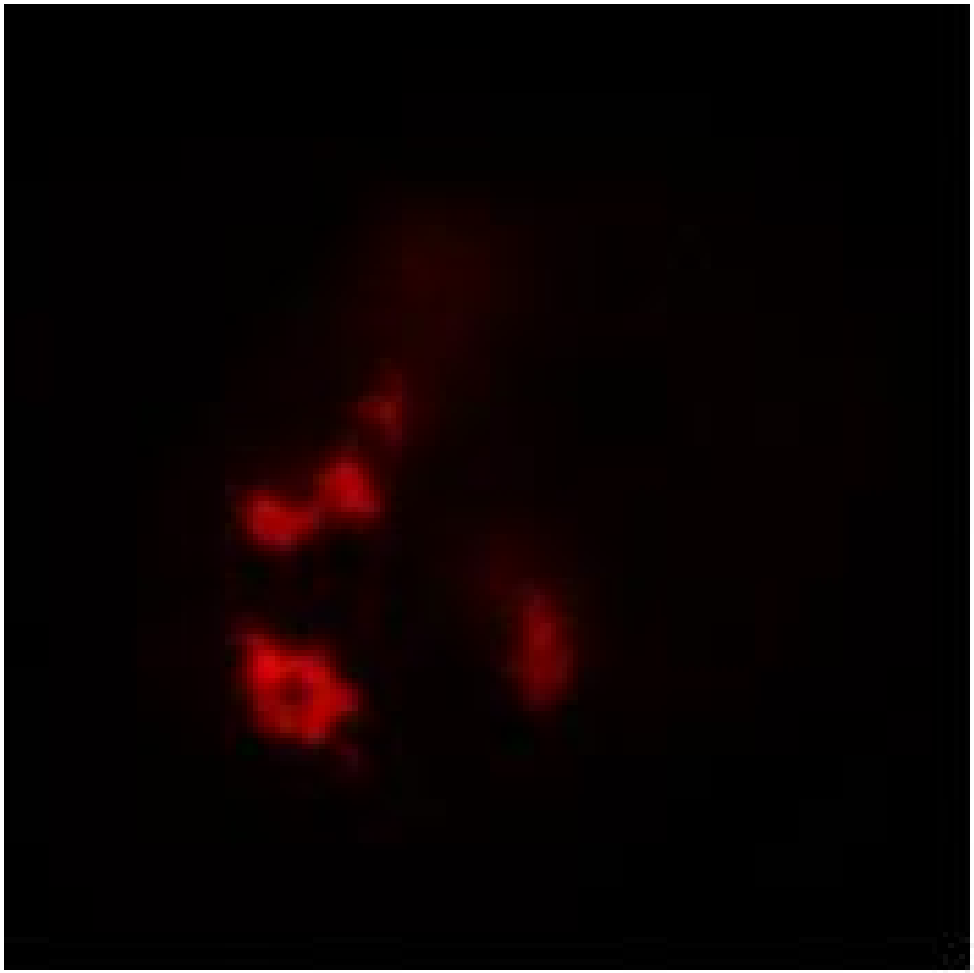
Slices: Nucleoli



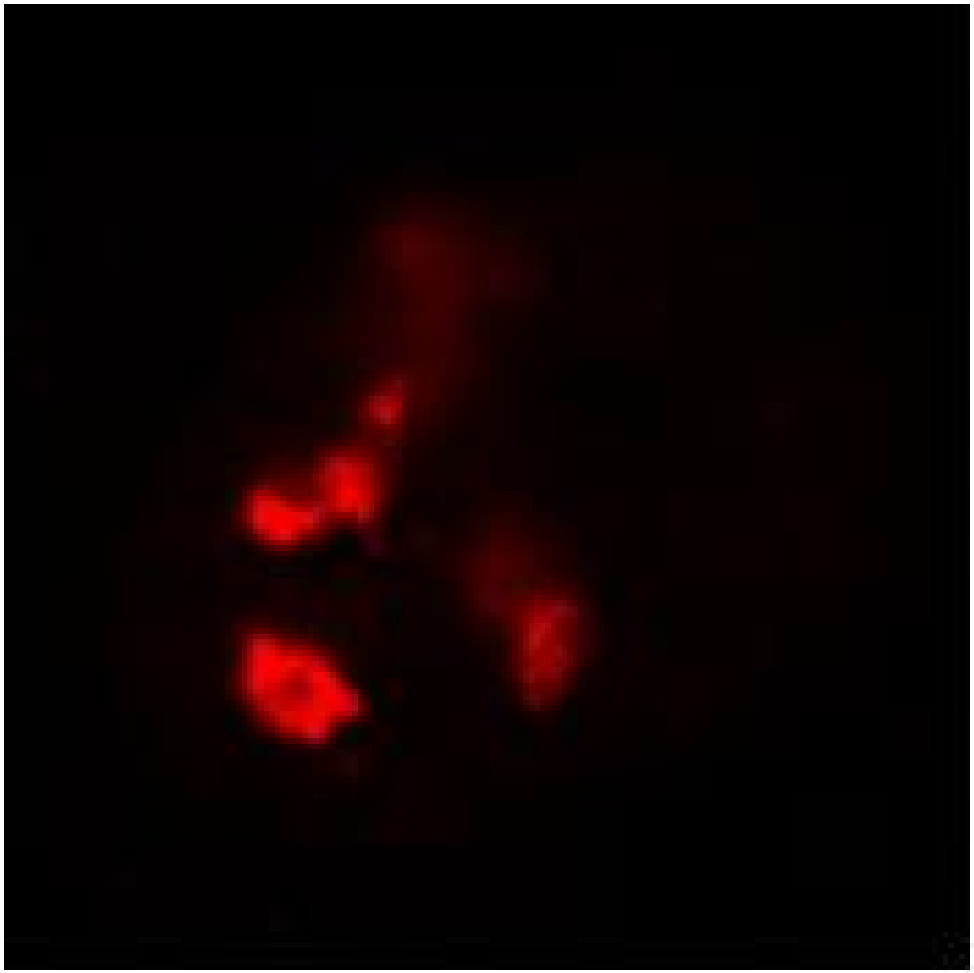
Slices: Nucleoli



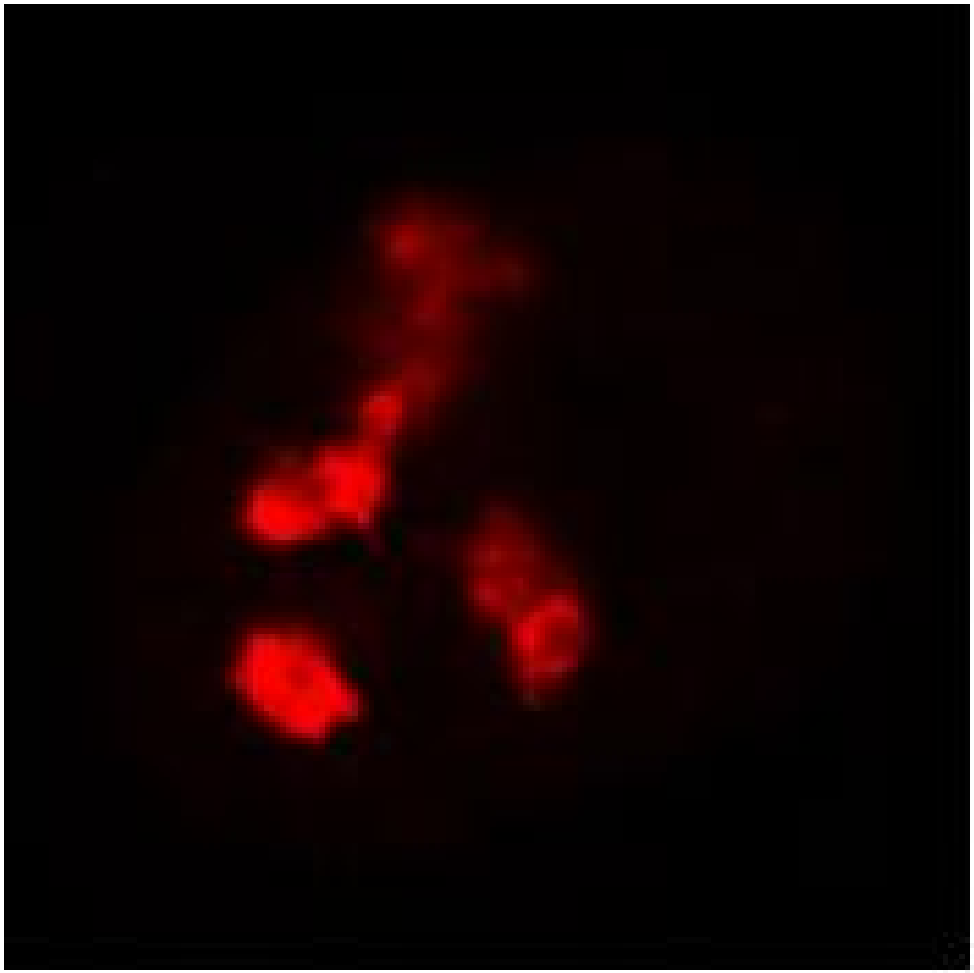
Slices: Nucleoli



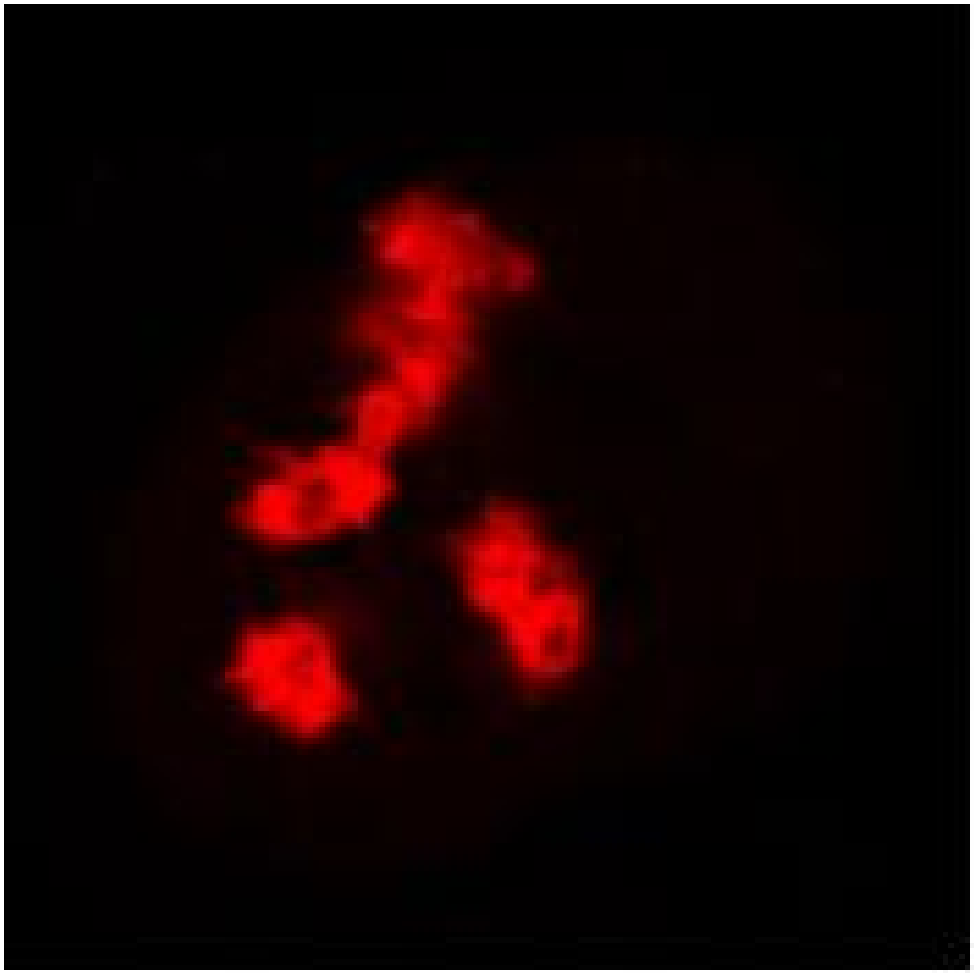
Slices: Nucleoli



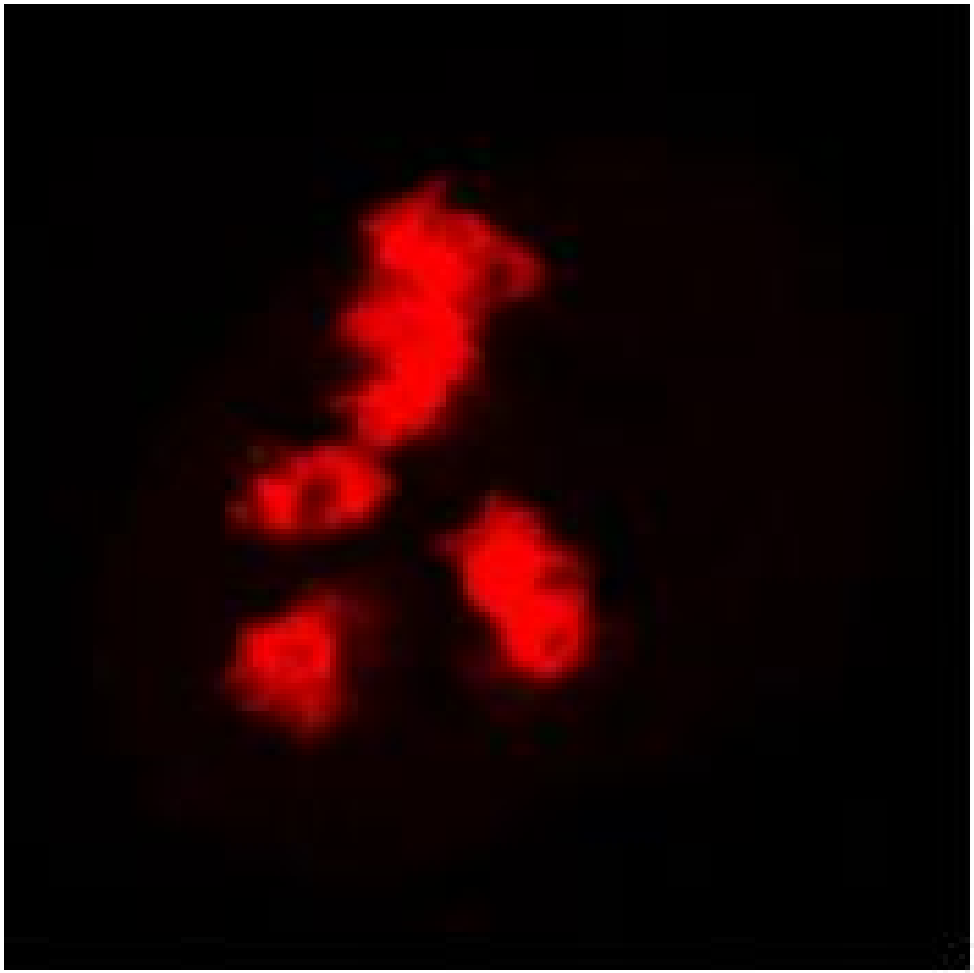
Slices: Nucleoli



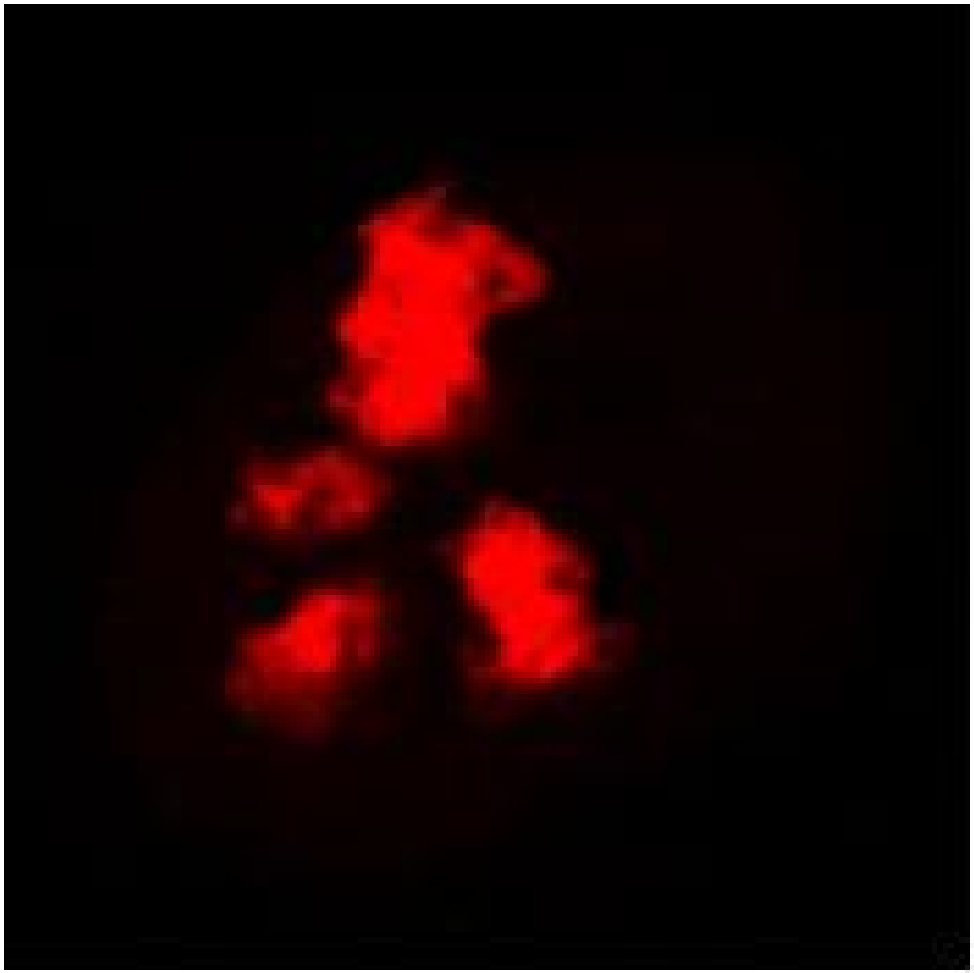
Slices: Nucleoli



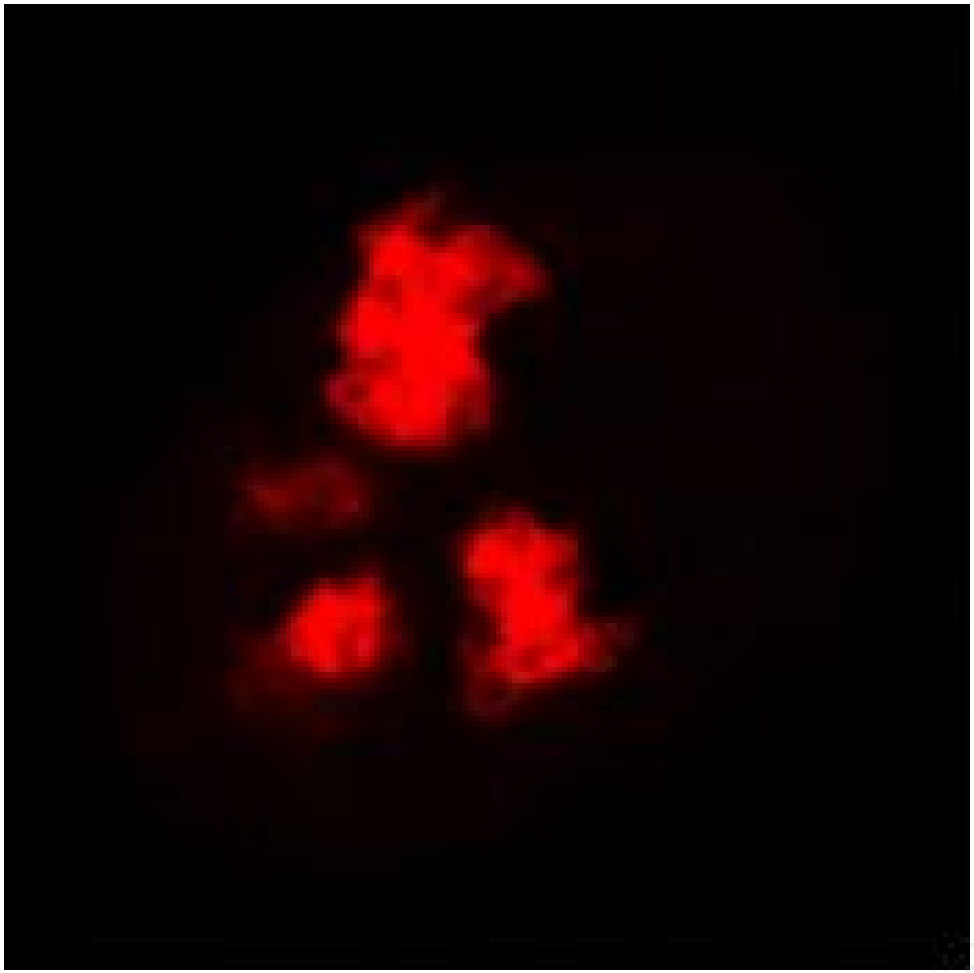
Slices: Nucleoli



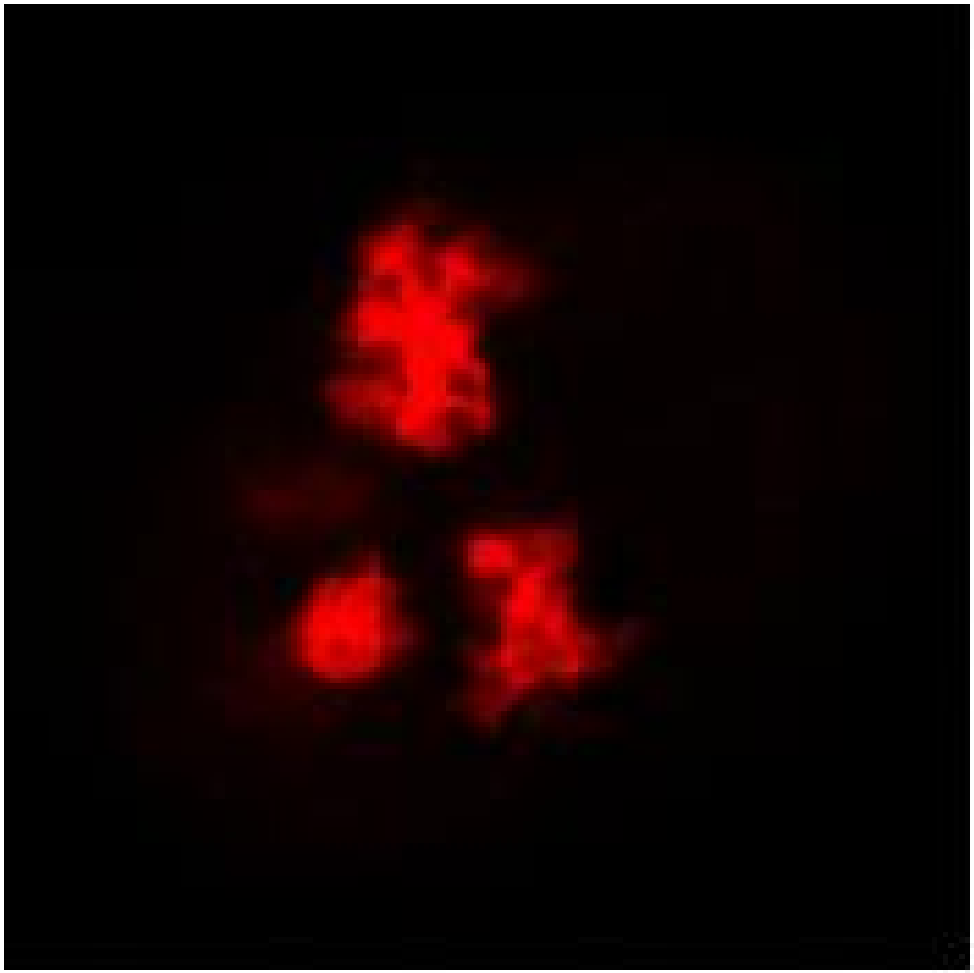
Slices: Nucleoli



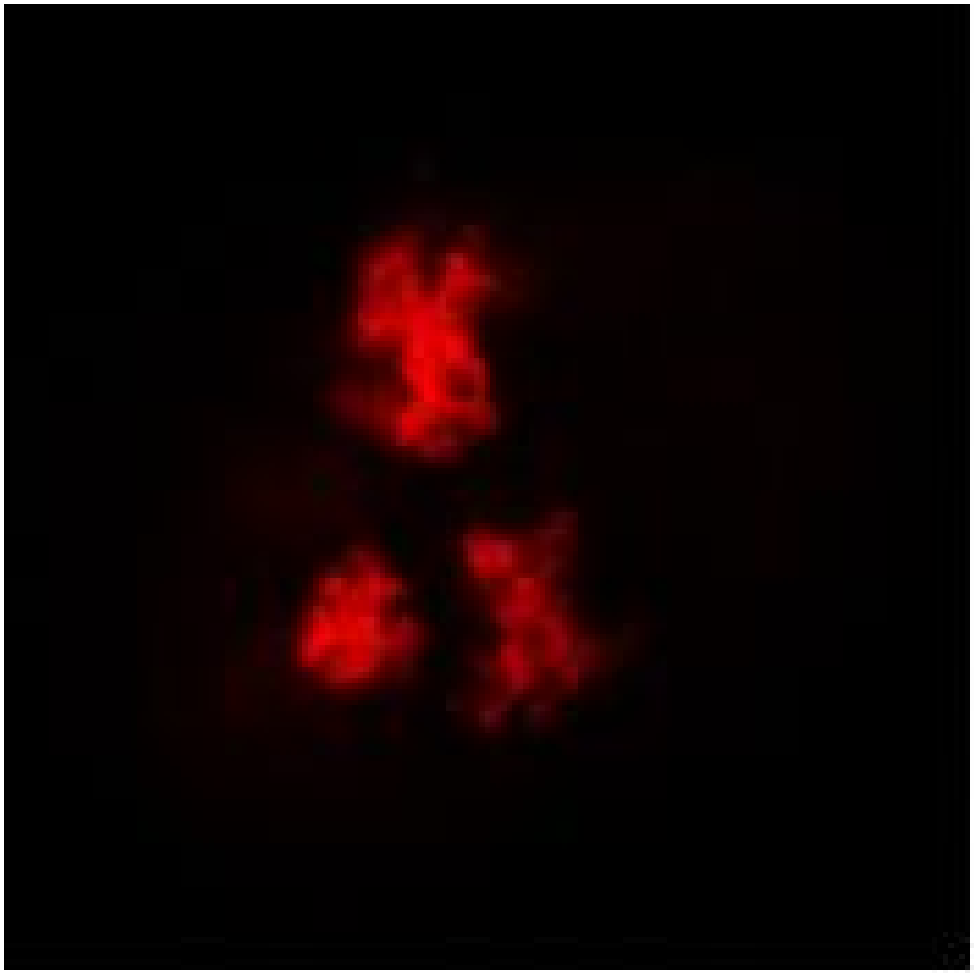
Slices: Nucleoli



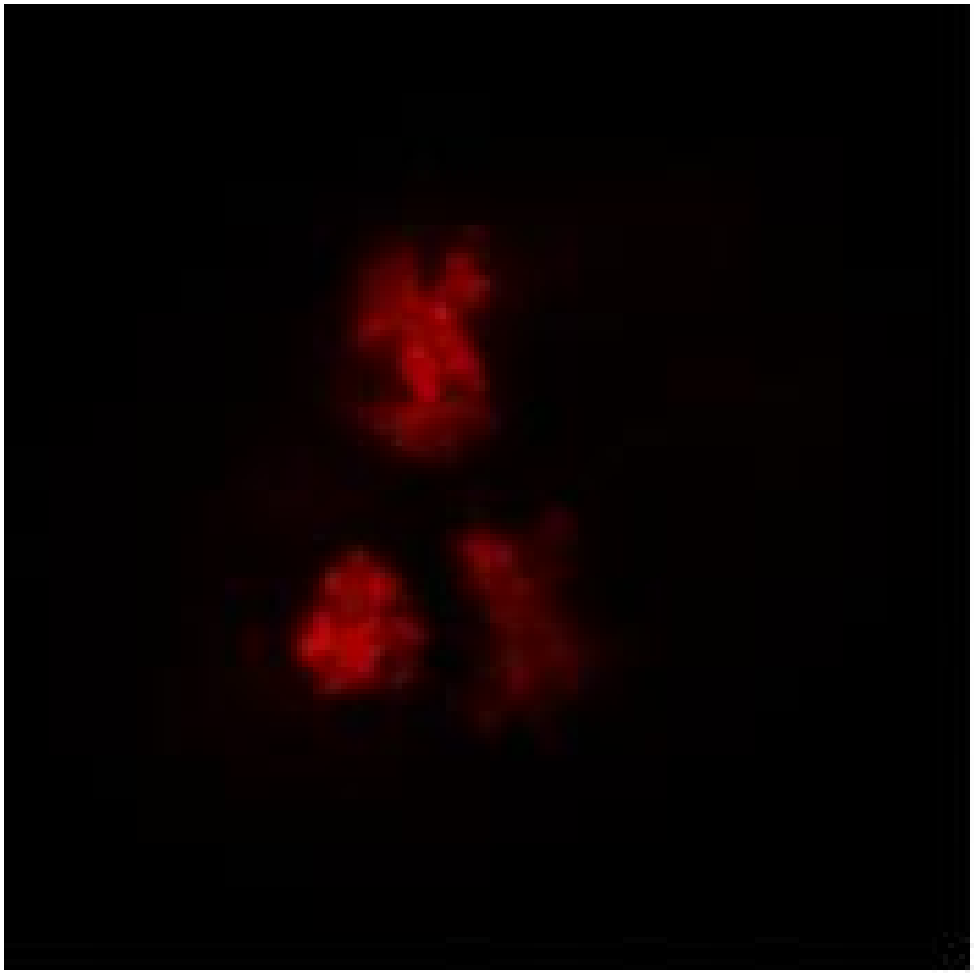
Slices: Nucleoli



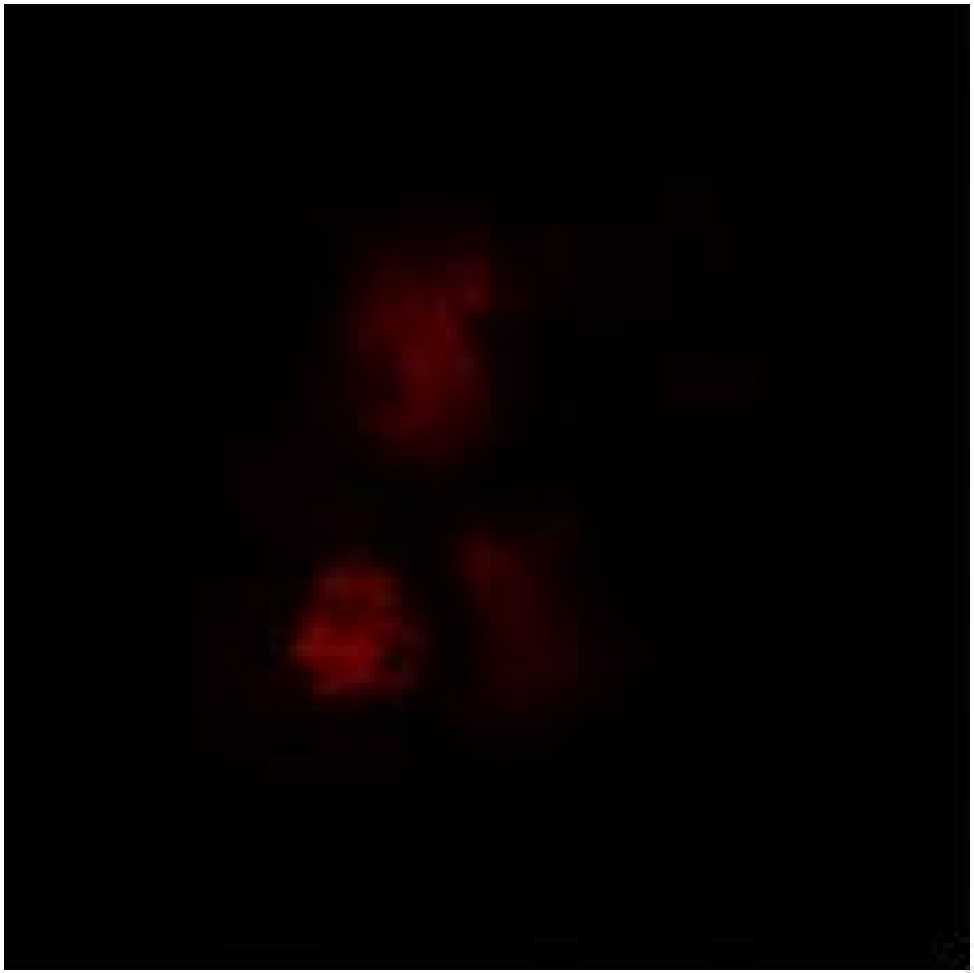
Slices: Nucleoli



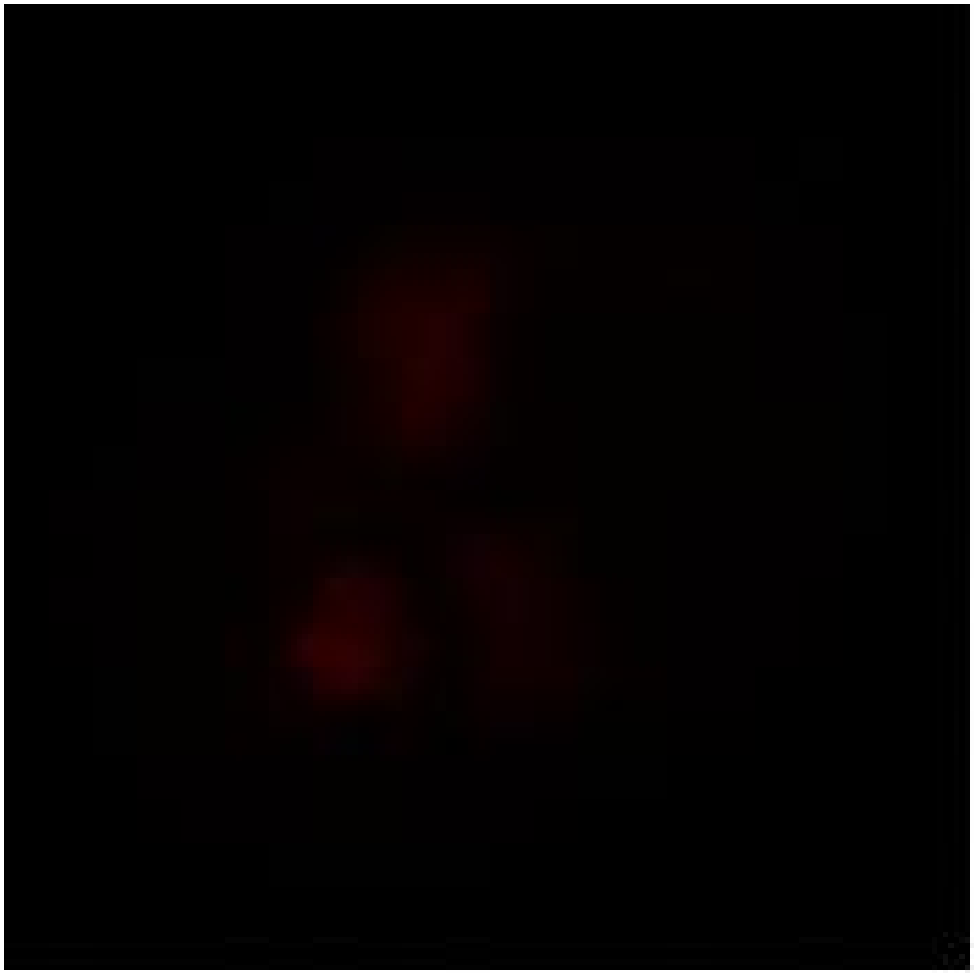
Slices: Nucleoli



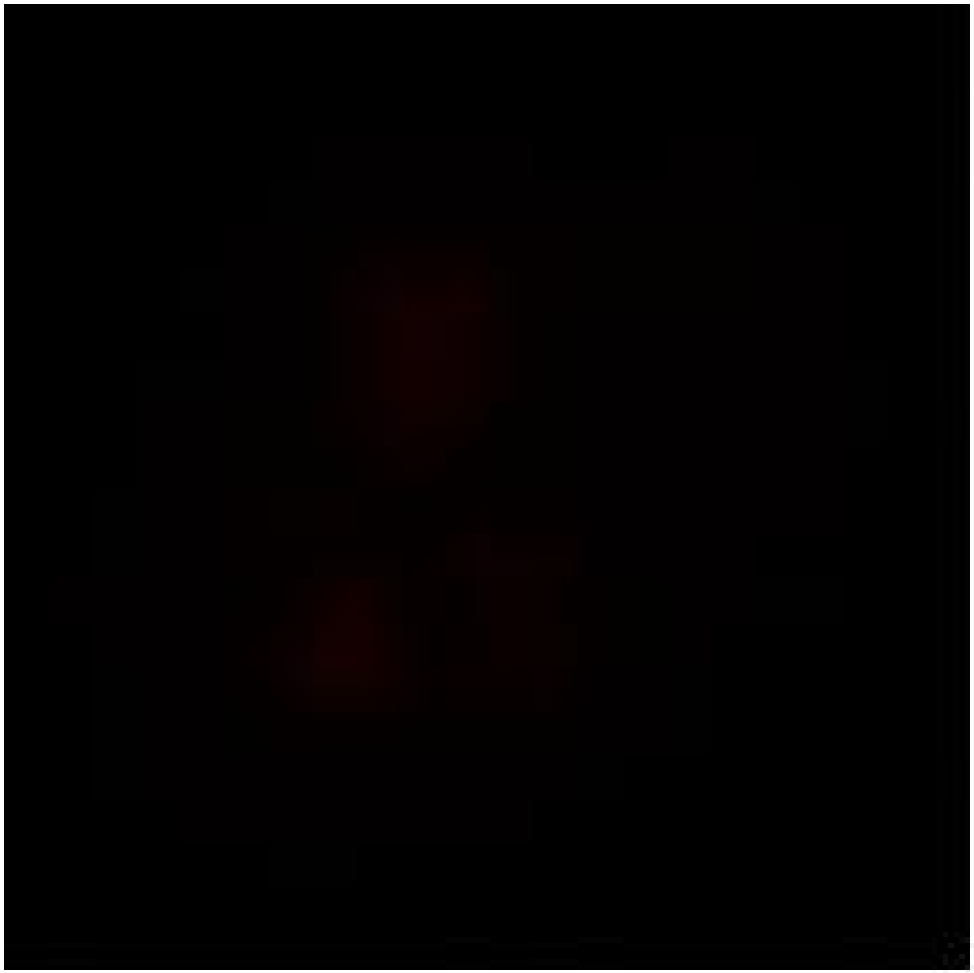
Slices: Nucleoli



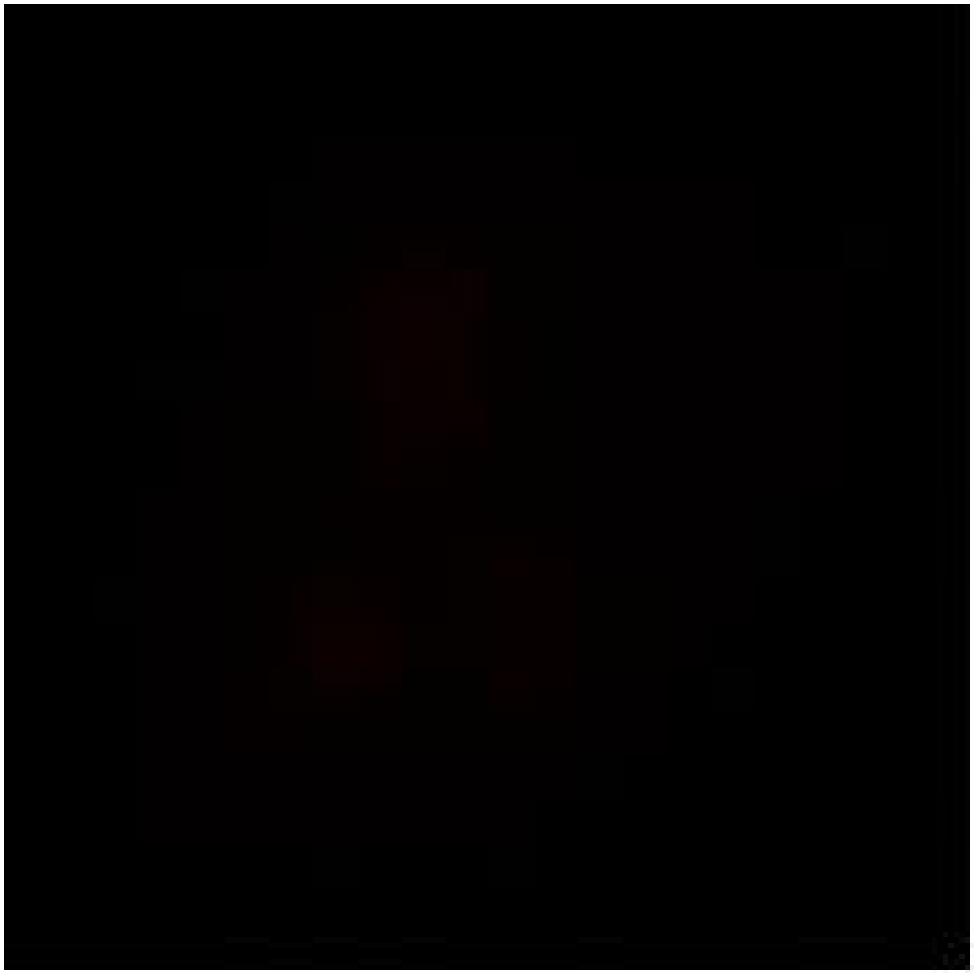
Slices: Nucleoli



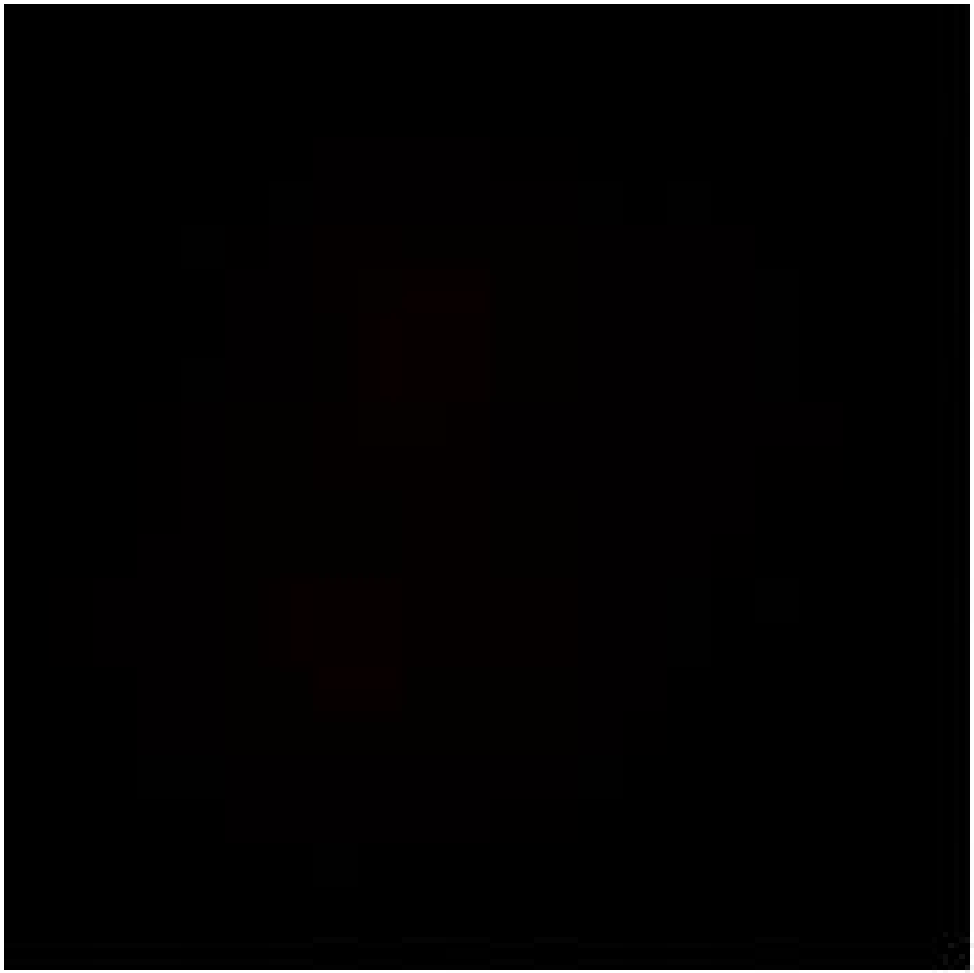
Slices: Nucleoli



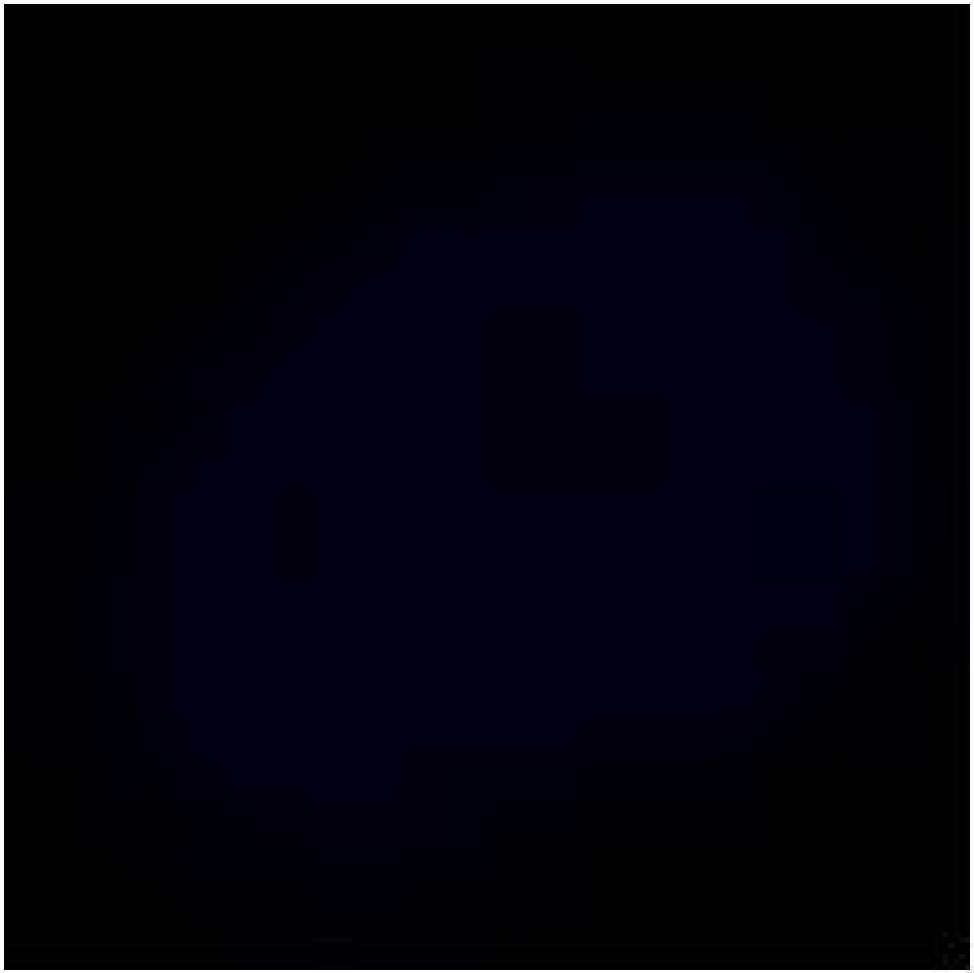
Slices: Nucleoli



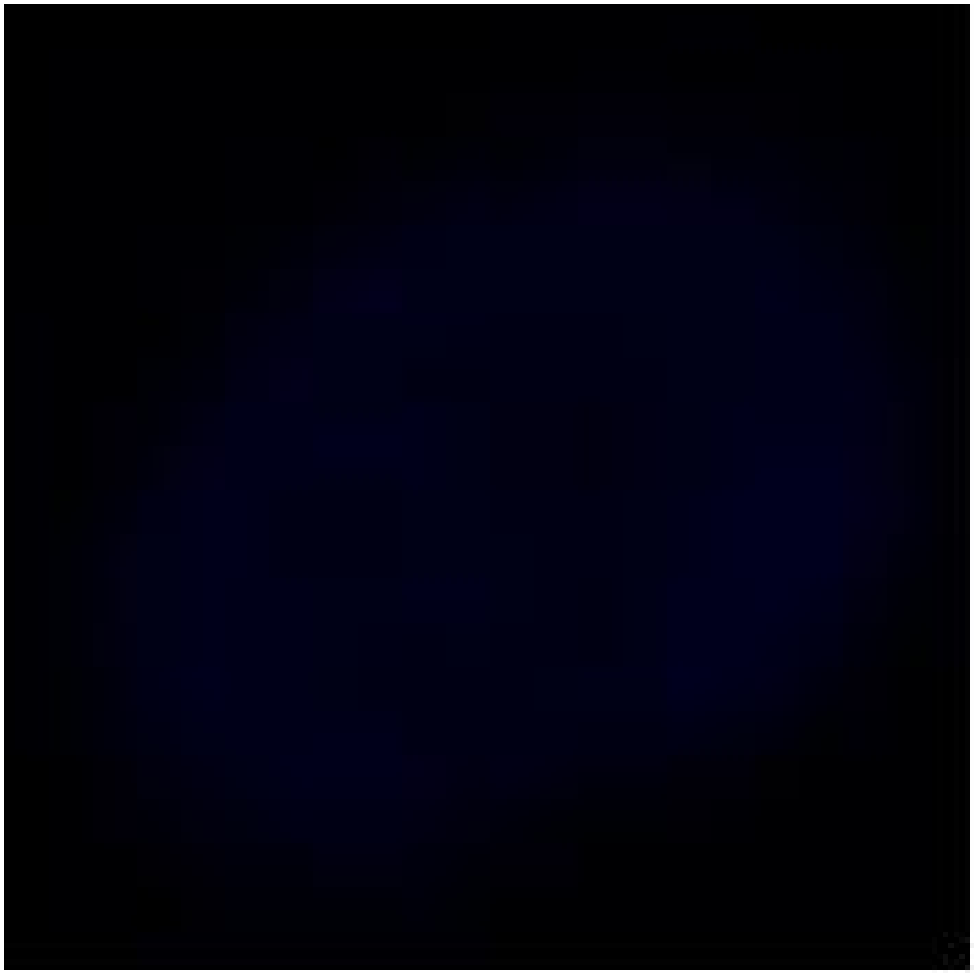
Slices: Nucleoli



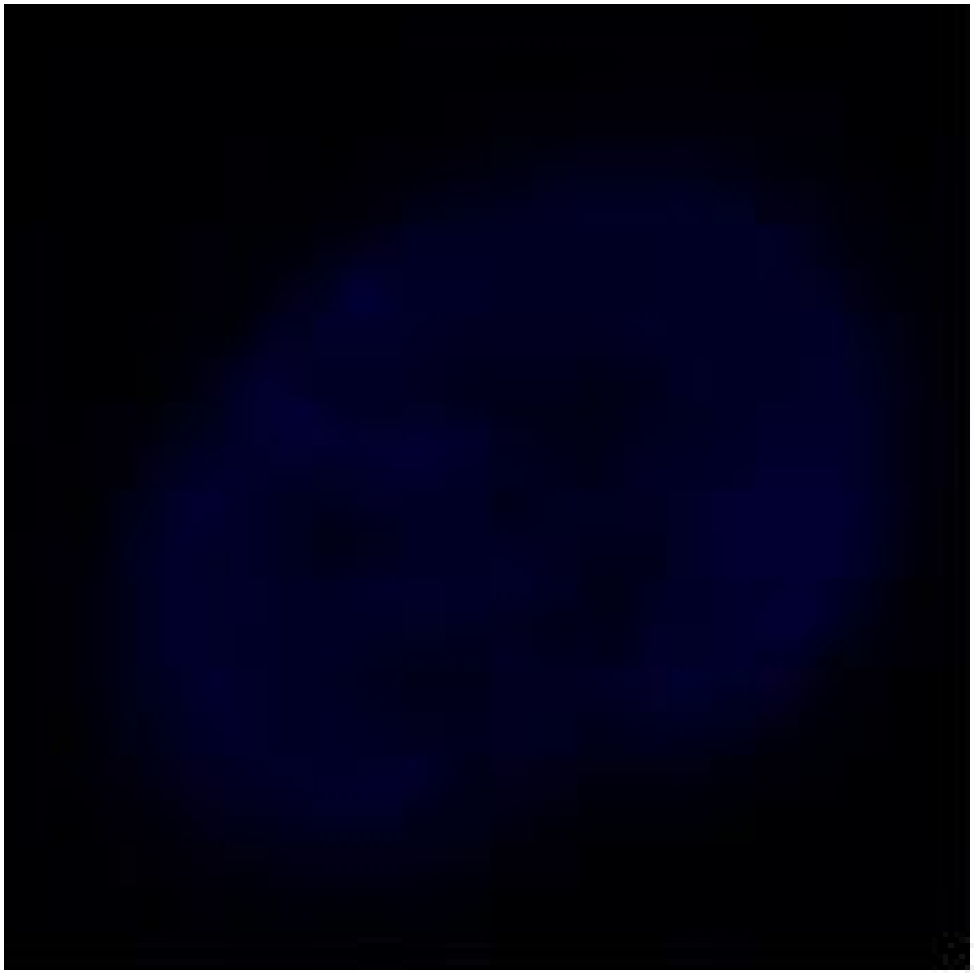
Slices: Nucleus



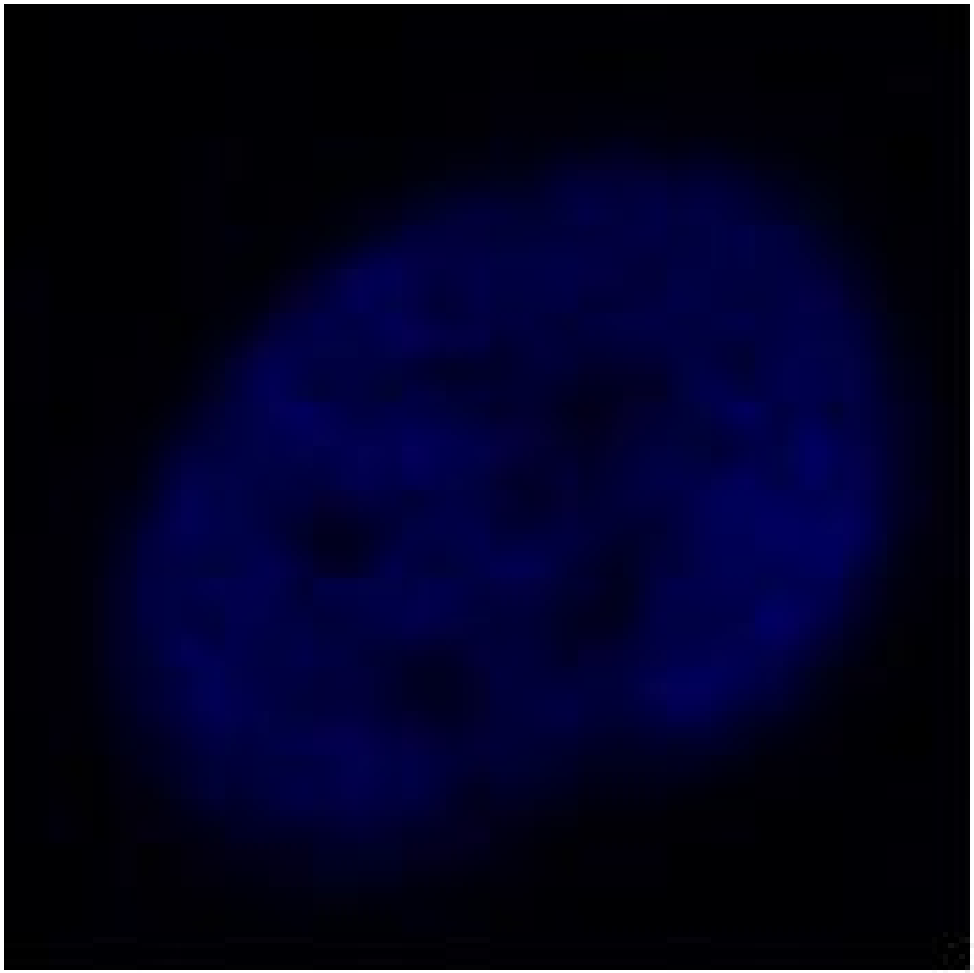
Slices: Nucleus



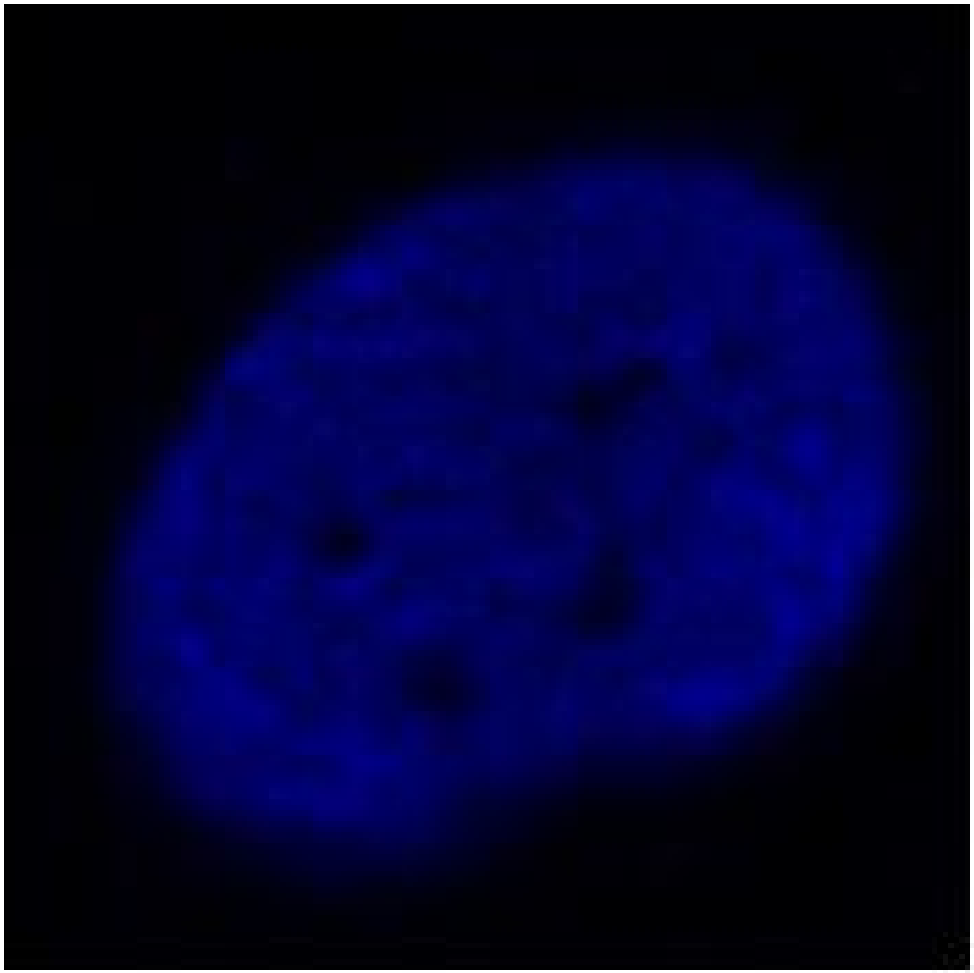
Slices: Nucleus



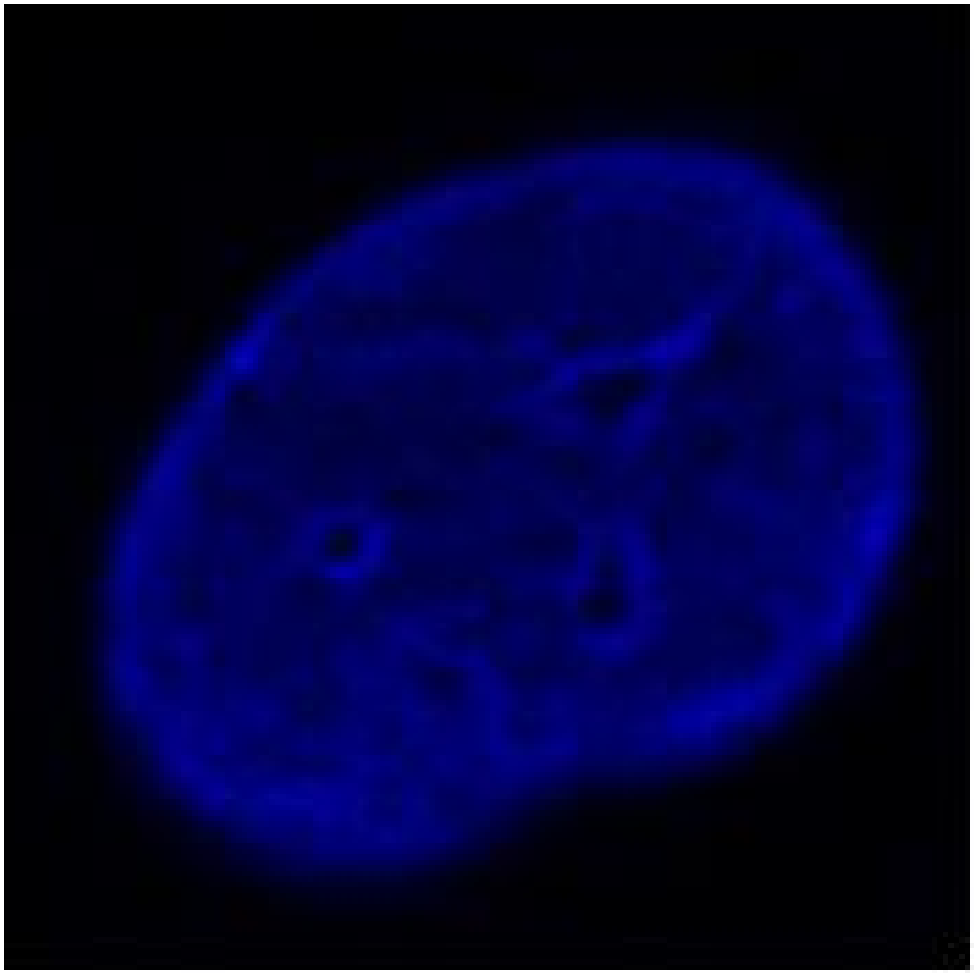
Slices: Nucleus



Slices: Nucleus

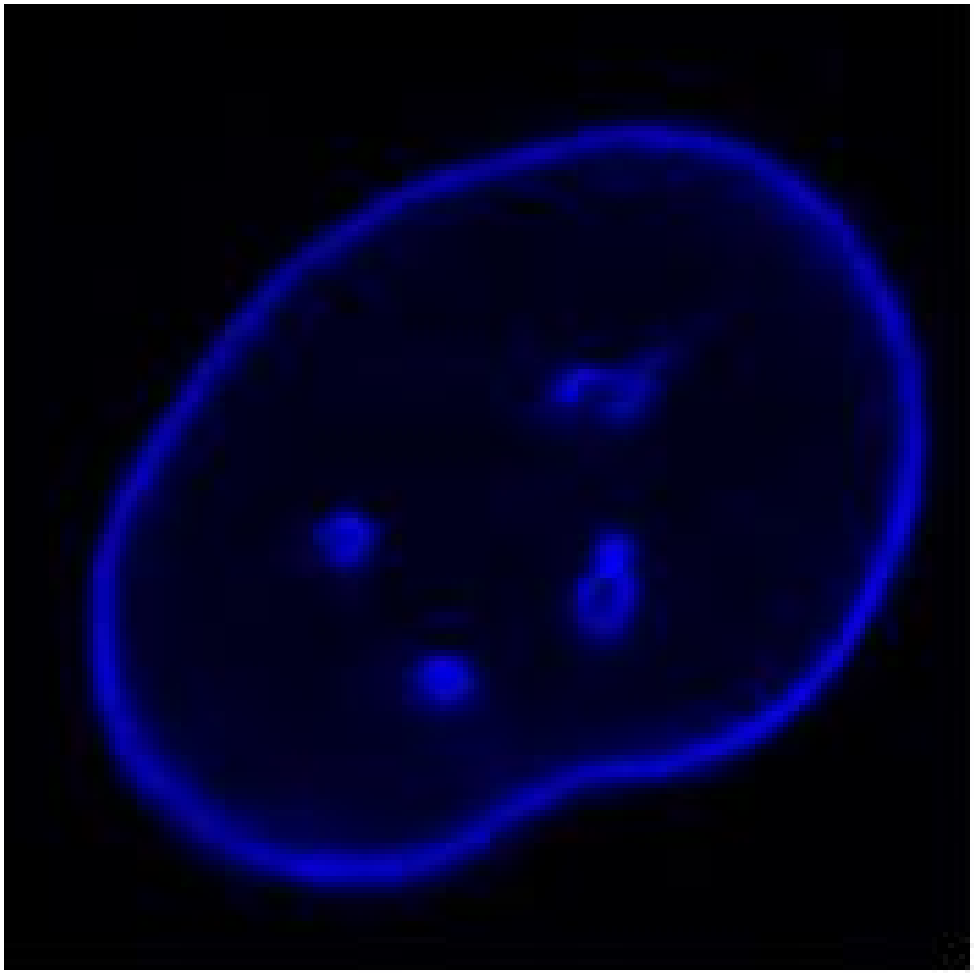


Slices: Nucleus

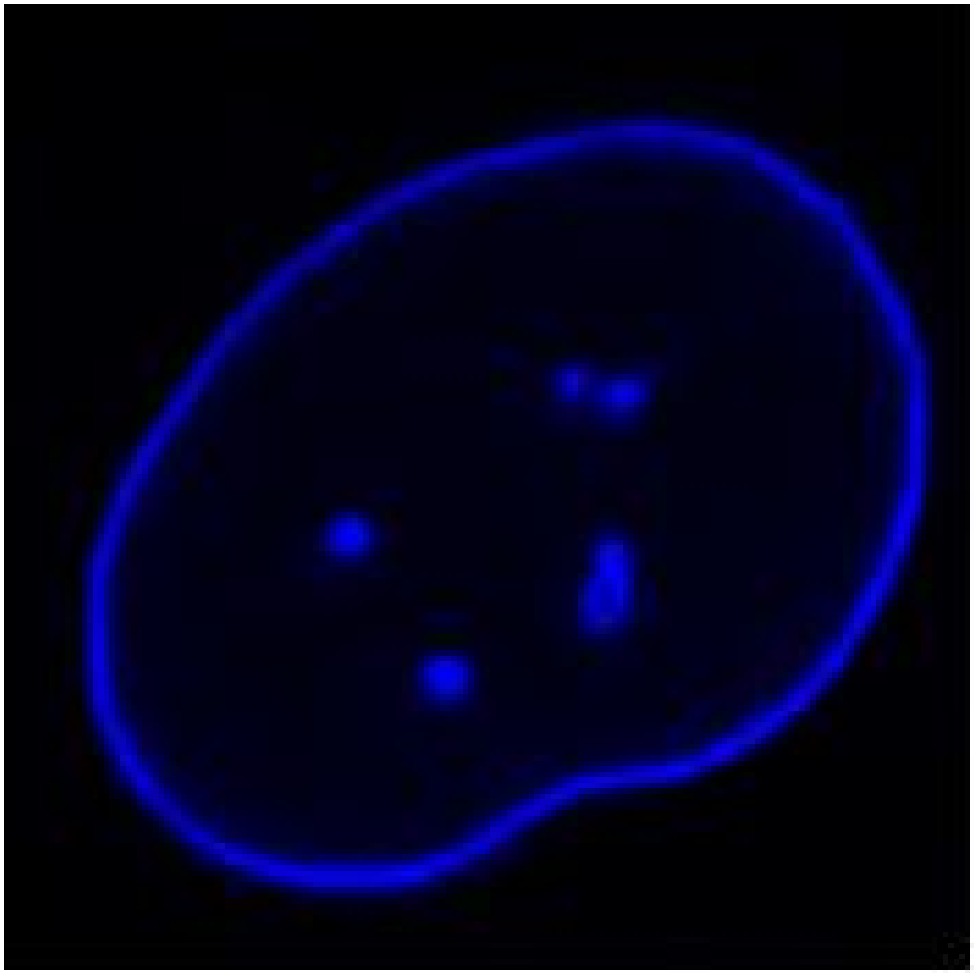


Slices: Nucleus

Slices: Nucleus

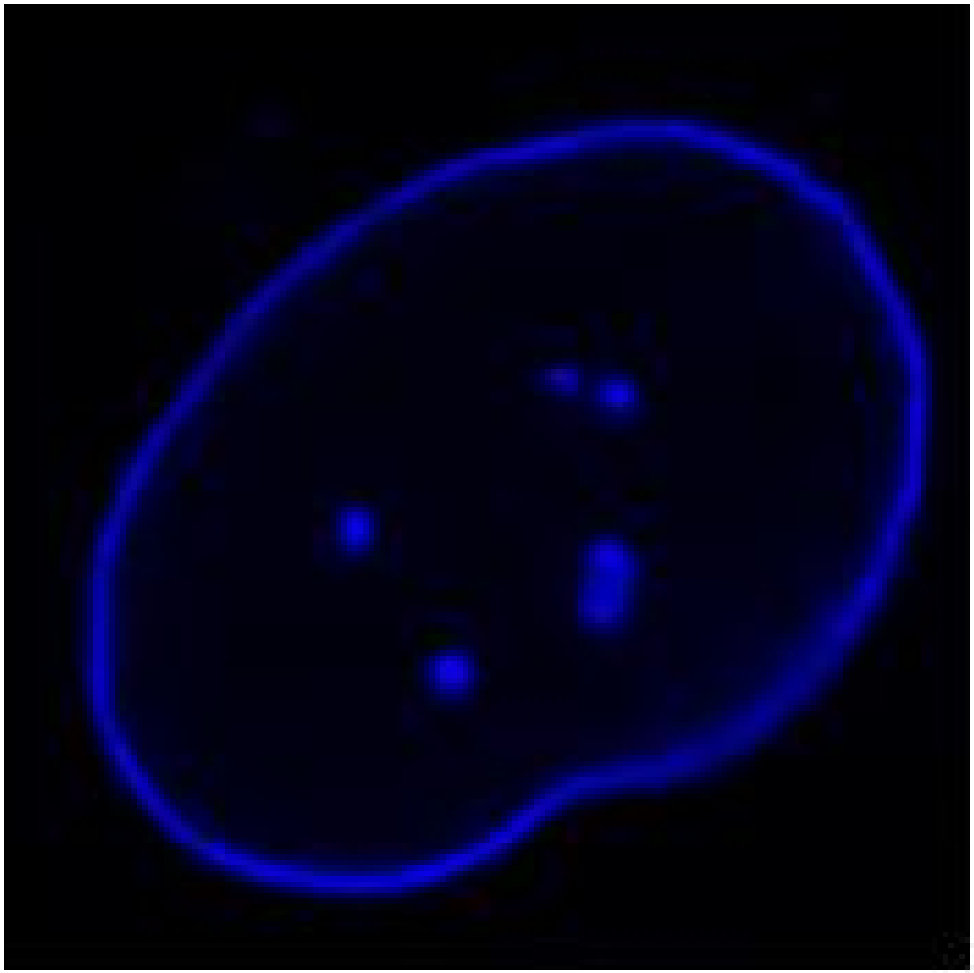


Slices: Nucleus

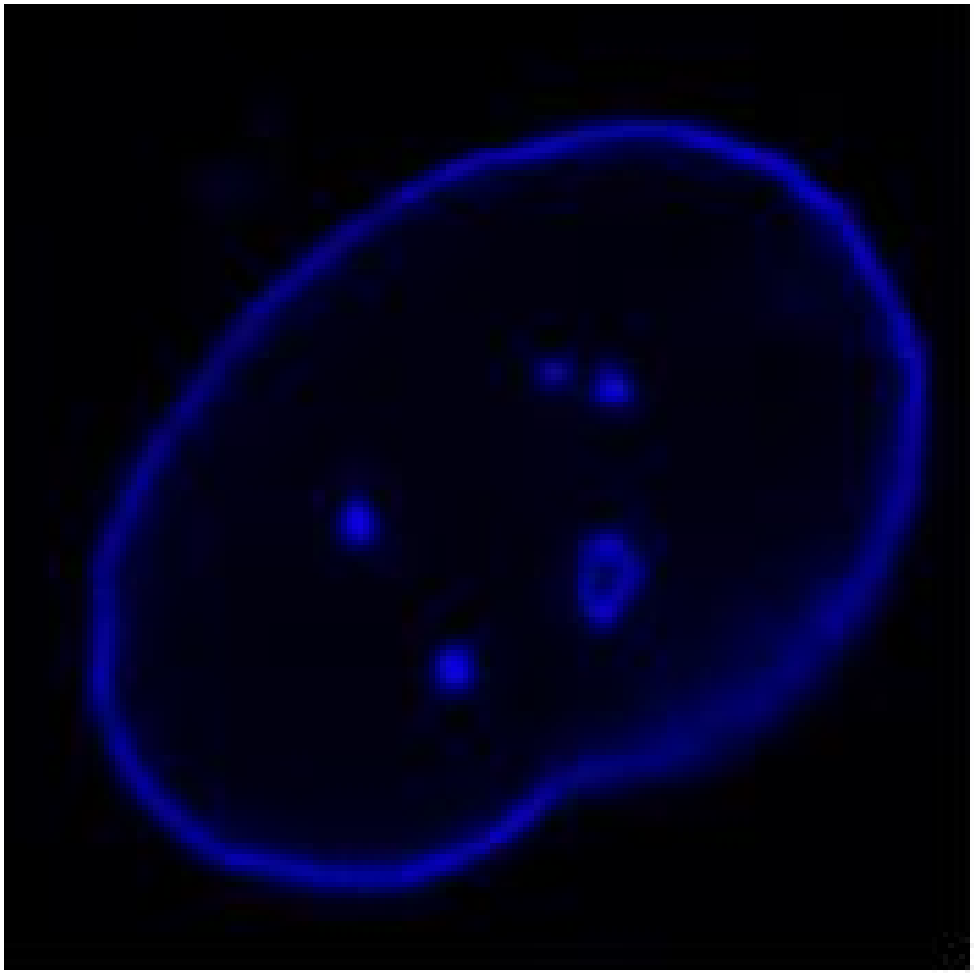


Slices: Nucleus

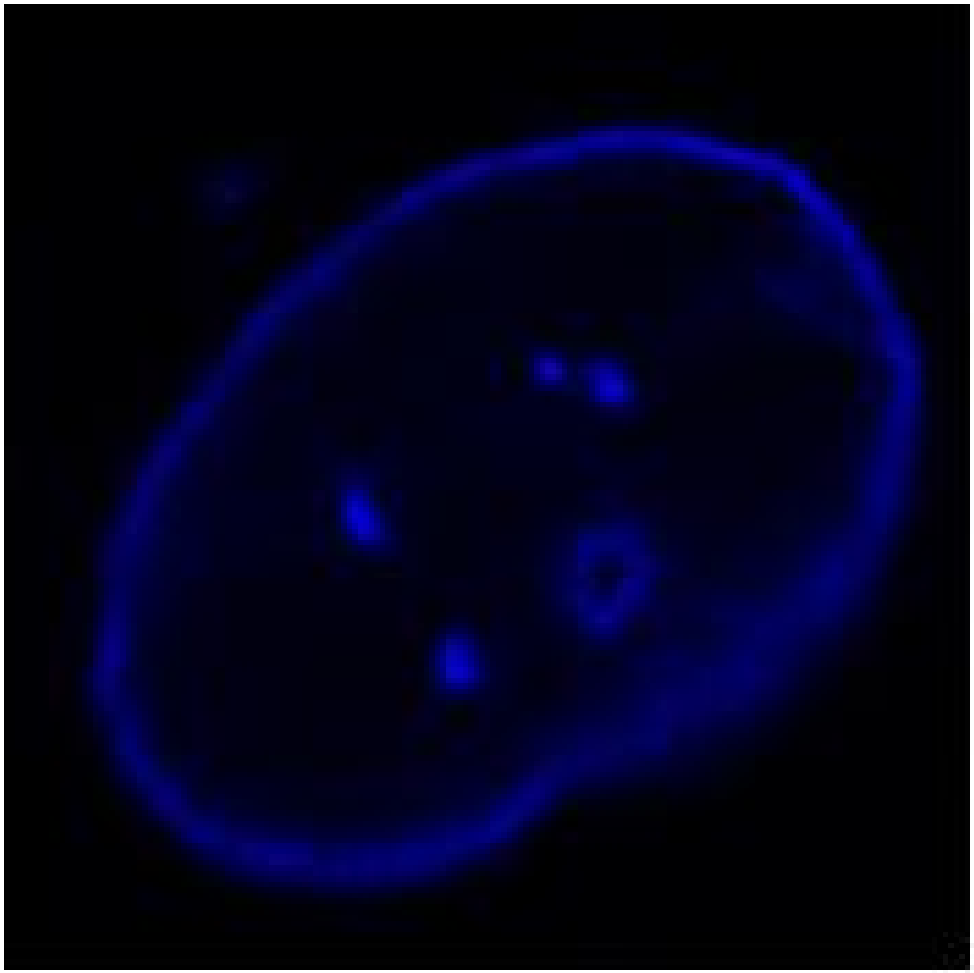
Slices: Nucleus



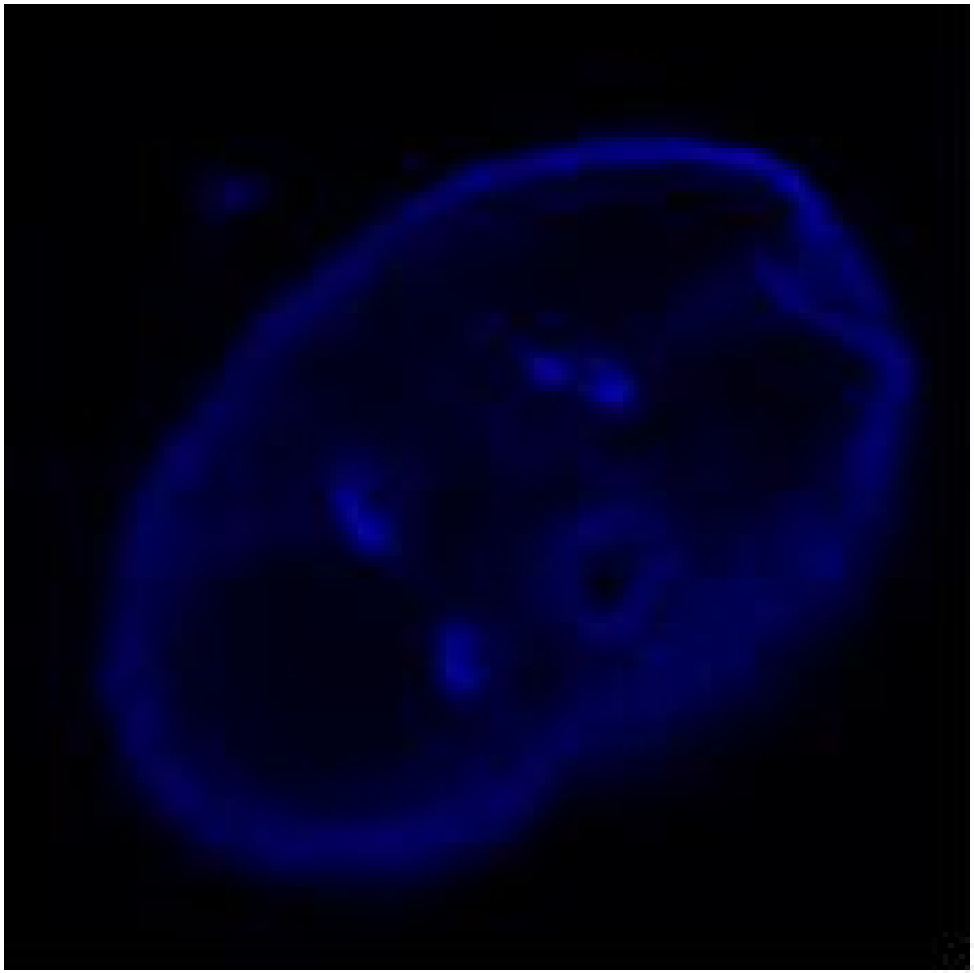
Slices: Nucleus



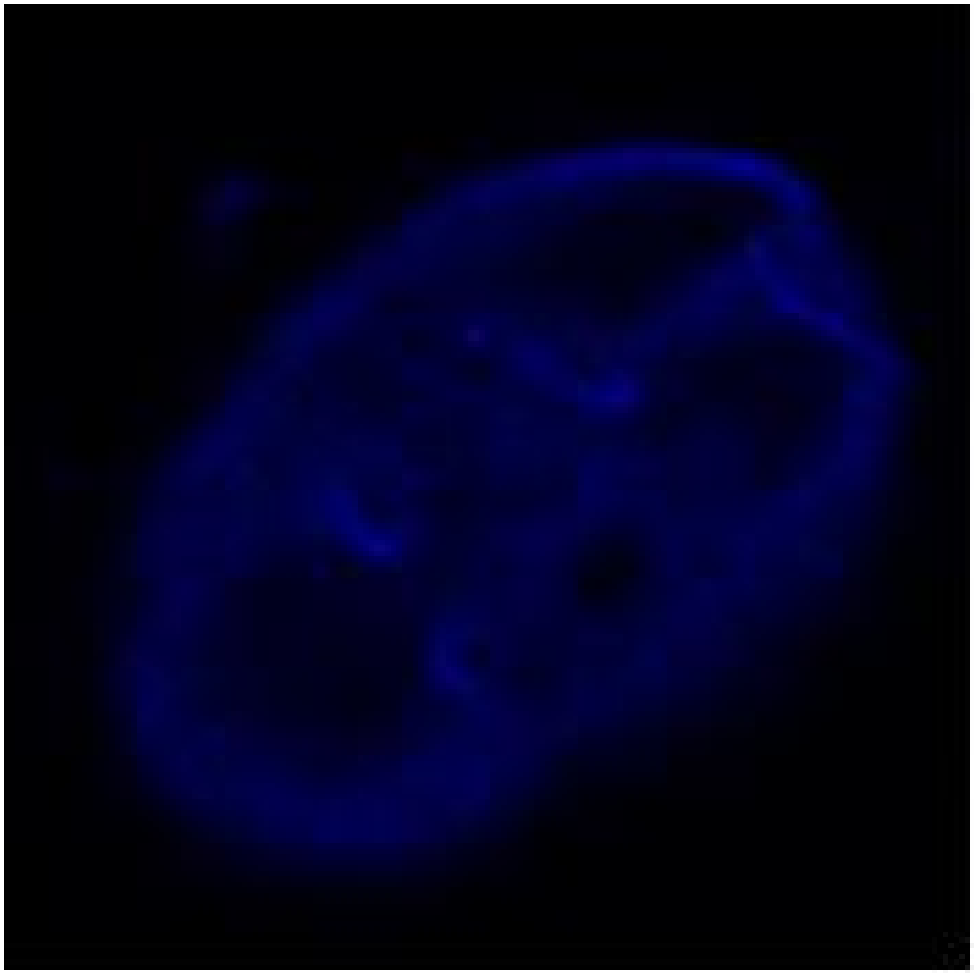
Slices: Nucleus



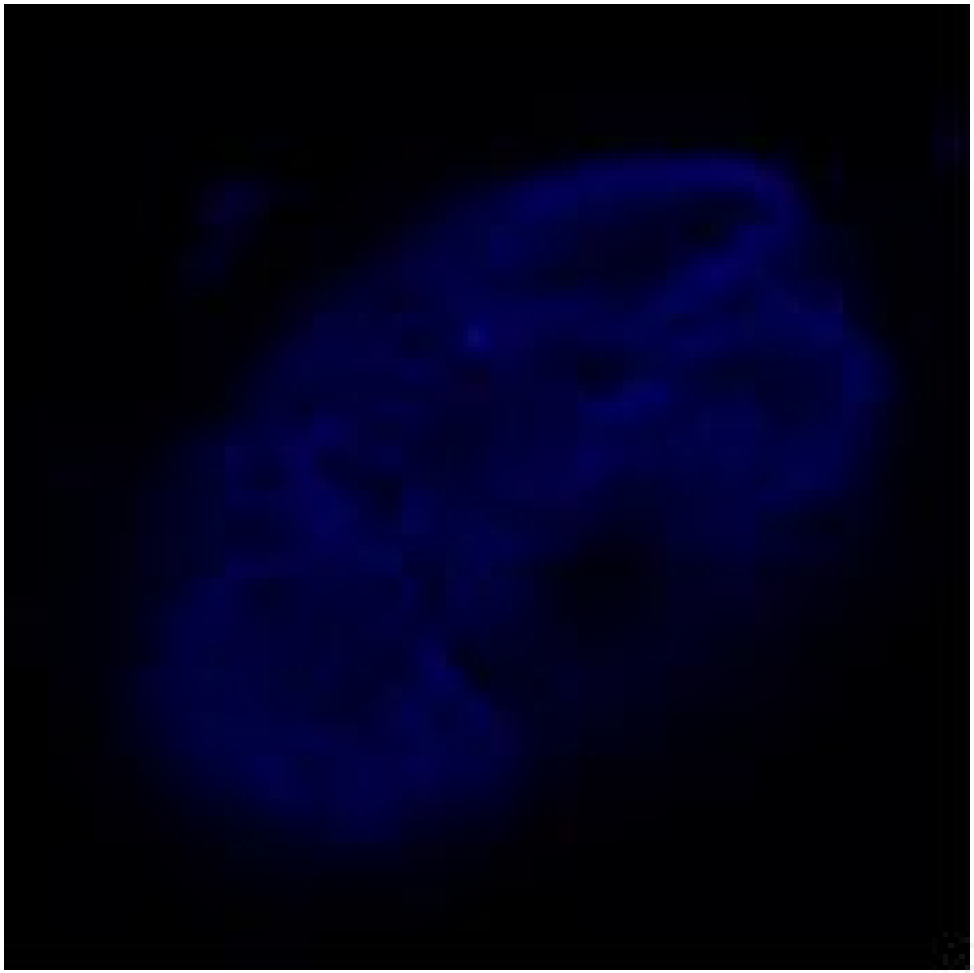
Slices: Nucleus



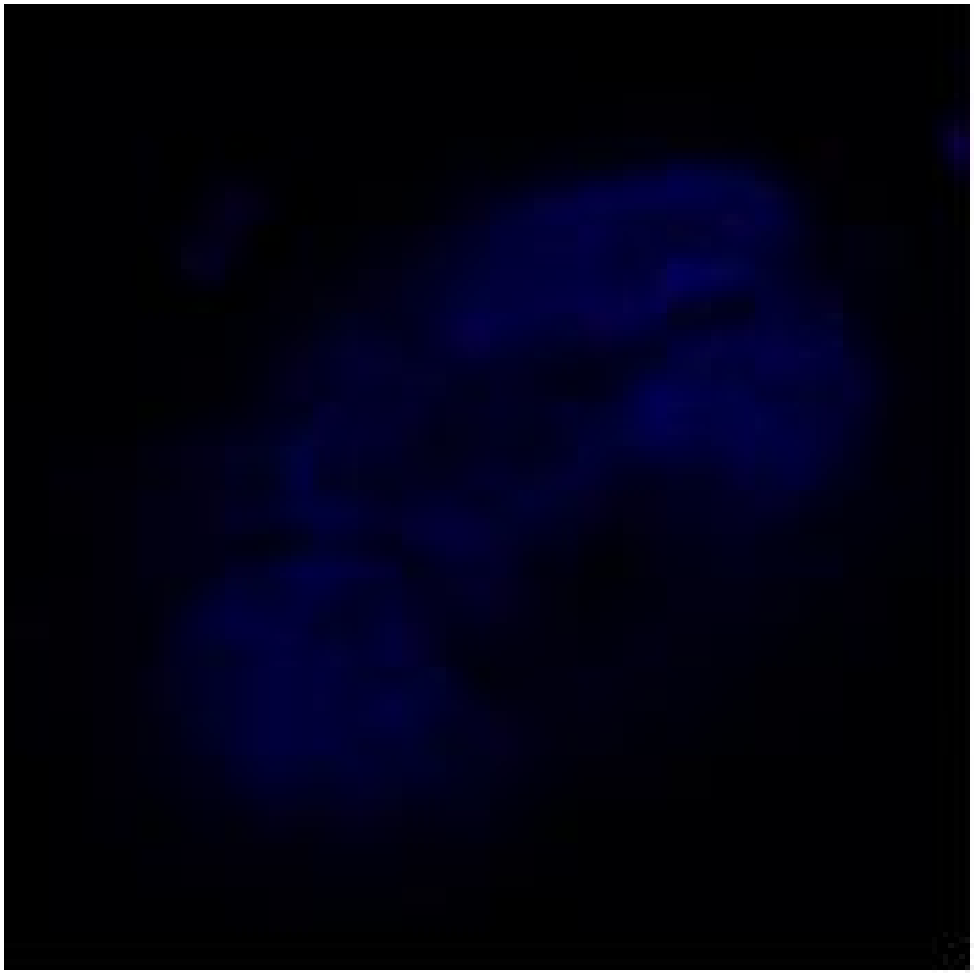
Slices: Nucleus



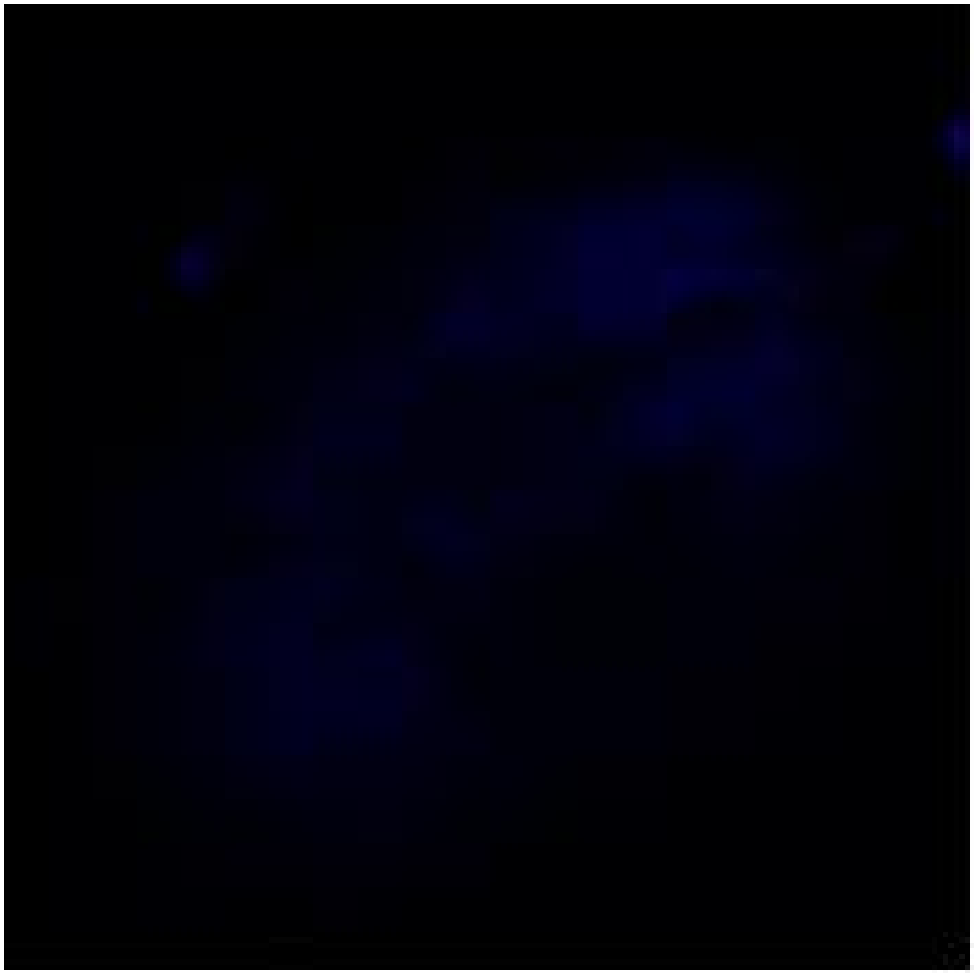
Slices: Nucleus



Slices: Nucleus

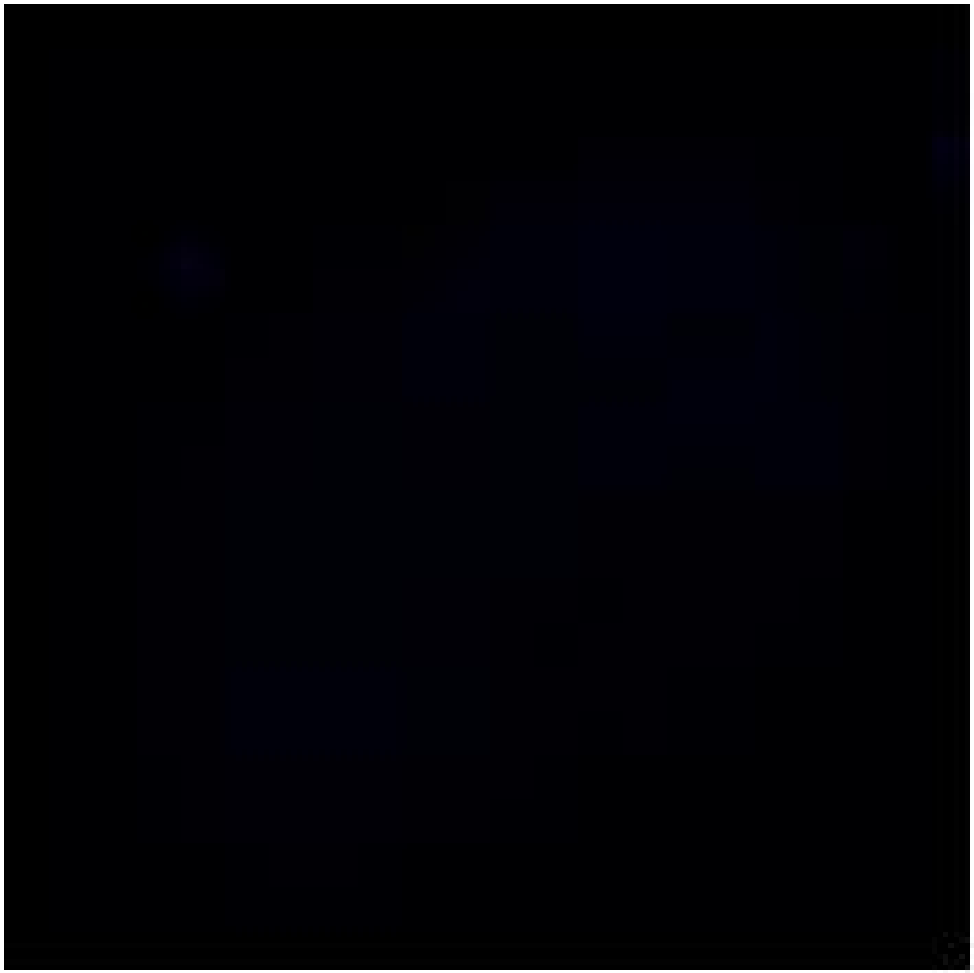


Slices: Nucleus

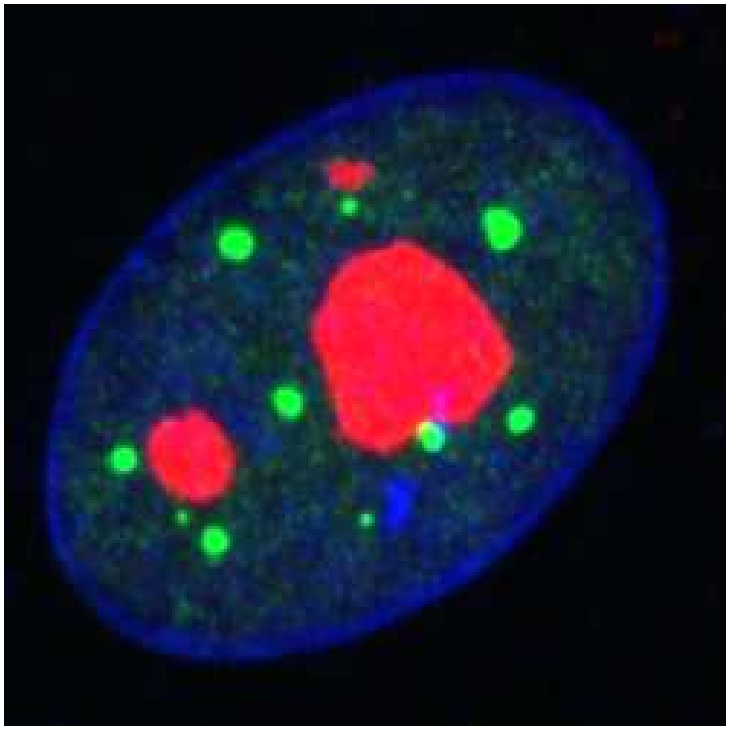


Slices: Nucleus

Slices: Nucleus



Combined Image



A Biological Hypothesis

Structures of similar function compartmentalize or locate adjacently, therefore, the comparative proximity of a particular nuclear compartment to PML bodies may indicate a functional relationship.

Many papers report the co-localization of PML with other nuclear compartments/objects; this observation is (almost without exception) made after visual inspection of images.

We seek to support/refute this hypothesis in a quantitative fashion using statistical means.

Data Analysis

Confocal microscopy data are available after careful preparation of cell examples.

Limitations of immunofluorescence procedure:

- At most three colour staining in the same image
  - PML/Nucleoli/Lamina
  - PML/Other Compartment/Lamina
  - Not possible to stain all compartments
- Cells different to synchronize
- $20 \times 250 \times 250$ voxel grid; surfaces must be interpolated
- Voxel allocation inferred by thresholding RGB levels; induces some uncertainty.

Data Analysis

Raw data processed into summary data;

- Locations in 3D of PML, other loci
- Ellipsoid approximation to nucleus
- PML volumes

Objective is to assess co-localization of objects within the nucleus.

Heroic Assumptions

Our idealized analysis framework will assume:

- cells synchronized
- population of replicate images independent/probabilistically identical
- two-object (PML/Body A) analysis in any one series of cell images is comparable with parallel two-object (PML/Body B) analysis in parallel series
- PML foci are points, not spheres.

Note: Magnification may vary from image to image; distances not absolutely comparable.

Statistical Analysis: Spatial Point Patterns (SPPs)

Hypothesis Tests based on Spatial Point Patterns (SPPs)

- PMLs only
  - Complete Spatial Randomness
  - Cluster Processes
- Two-object images
  - Cross-object K
  - Doubly-stochastic processes

Note: majority of models for SPPs for processes on unbounded, contiguous regions in $\mathbb{R}^2$; extensions to bounded, non-contiguous regions in $\mathbb{R}^3$ represent considerable methodological challenge.

Issues

- Edge Correction
  - Common problem in SPPs; correction for observation on bounded field of vision; need to use edge-correction adjustments to standard statistical summaries.
  - Here, field of vision is fixed and bounded; however, most objects are excluded from nucleus periphery.
- What is “Distance” ?
  - Interior of ellipsoid
  - Conformation of nucleus is irregular
  - Nucleoli are large exclusion regions

Statistical Analysis

Hypothesis Tests based on Distance

- Comparison of Nearest Neighbour (NN) distance distributions
  - t-test
  - Mann-Whitney Wilcoxon
  - Kolmogorov-Smirnov
- NN log-odds modification for two-object problems
  - compute log-odds on NN being same/different object
  - standardize for numbers of objects of different types
  - reduces to binary GLM for NN counts

Randomization Tests

Distributional assumptions not appropriate (for NN distance comparisons); forced to use randomization (or permutation) methods.

1. Propose test statistic, T, compute for observed point pattern, $t^\star$.
2. For $i = 1, \ldots, B$ replications;
   (i) Randomly (re)sample "new" data set from observed data by bootstrap/randomization/permutation
   (ii) Compute new test statistic, t_i, evaluating T for new data set.
3. Compare $t^\star$ with null distribution formed by $t_i, i = 1, \ldots, B$ to evaluate statistical significance.

Note: need to combine information from multiple cell images. This is a non-standard problem in SPPs.

Respecting the by-image nature of the point pattern data within the randomization is necessary and not necessarily straightforward.

We can utilize more formal inferential methods in a model-based setting;

LIKELIHOOD

This gives an optimal framework for combination of information.

Likelihood

Many parametric models exist for single/paired point patterns:

- Parameterized via intensity process or measures of spatial dependence
- Can represent attraction/repulsion models
- Likelihood usually available, sometimes via missing data representation
- Inference often requires advanced computation (say MCMC)

This is becoming fairly routine in 2D, but virtually unknown in 3D.

Some Results

Extensive analysis of

- PML only
- PML vs other loci
- CBP vs other loci

Results in different cell lines, for different organisms.

Summary:

- Reject Hypothesis of CSR !
- Evidence for spatial association between PML and several other compartments !

Future Experiments

- Advanced simulation-based testing
- Classification
- Cell Synchronization
- Dynamic Tracking
- RNAi peturbation experiments
- Different (whole image) summaries and comparison

Most require urther careful benchwork

The apparent intellectual barrier ...

> *It is clear that PML bodies lie in close proximity to other nuclear compartments/genomic loci; that they do so is a common experimental observation. Furthermore, it is conceptually clear that as there is some hierarchy of organisation within which PML bodies operate, this close proximity is hardly surprising. Hence your statistical analyses, although correct, are biologically worthless ...*

Reviewers, *passim*.

This is easy for us to refute; observation is not quantification or comparison, biological folklore is not evidence. This is where we need to convince the Reviewers.

Acknowledgements

Imperial College Division of Biomedical Sciences

- Elizabeth Batty
- Suhail Islam
- Paul Freemont
- Carol Shiels

Imperial College Department of Mathematics

- Niall Adams
- Philip Umande