

Supplementary information

Recoverable temporal and spectral information in public human macrophage imaging

A reproducible case audit

Elkin Navarro Quiroz

Universidad Simón Bolívar, Barranquilla, Colombia

S1 Scope and analytical units

This supplement accompanies an exploratory secondary audit of four published human macrophage videos and one fixed-cell stimulated Raman scattering (SRS) field. Files and fields are the descriptive units. Frames, pixels and tiles are dependent observations within them; five files from three studies do not represent five donors or five biological replicates. The original biological acquisitions belong to the cited source investigators.

The temporal repetition analysis concerns the representation of one PLOS movie. The repeated-coordinate analysis concerns two planes of one fixed SRS field. Agreement between algorithms, readers or repeated executions assesses computation on those inputs. No tracking algorithm, phenotype classifier, molecular-decomposition model or biological event adjudication was evaluated.

S2 Selection record and chronology

The focused search began on 21 October 2026 in Colombia. Its preserved routes include exact title/DOI searches, publisher and PMC/Europe PMC text, BioStudies metadata and file inventories, Zenodo metadata, supplementary material and OpenAlex availability metadata. Exact search-engine query strings, result counts and times for every search were not retained. They cannot be reconstructed as a systematic search. The source ledger records what was examined, access limits and admission decisions; download records retain actual retrieval times in UTC.

Eligibility required a retrievable image/video, a traceable human macrophage source and sufficient modality information. Selection was purposive and sought complementary representations relevant to a later cell-history question. It did not presuppose that the original sources were designed as tracking benchmarks. A larger set, a paired native/exported acquisition or a reference-annotated benchmark might support other questions; none was added to the present empirical set.

Table S1. Resource selection and role. The original source ledger is retained alongside this retrospective summary. Bibliographic precedents are not numerical inputs.

Resource	Inspected material	Decision and limit
Canton 2016, DOI 10.1038/ncomms11284	Article, movie captions and complete movies 2/4/6	Three human video representations admitted; not three independent studies
Santulli-Marotto 2015, DOI 10.1371/journal.pone.0145078	Article, S1 caption and complete movie	Human efferocytosis rendering admitted; no new cellular event labels
Szkalitsy 2025, DOI 10.1038/s44318-025-00423-2; S-BIAD1584	Source methods, API, inventory, README and panel-3D TIFF	One fixed human SRS field admitted
Stiebing 2017, DOI 10.1002/jbio.201600279	Publisher abstract and bibliographic metadata	Reported live-human Raman/SRS precedent; complete methods and raw data not audited
Sapudom 2017, DOI 10.1038/s41598-017-14458-x	Primary article and availability statement	Human tracking precedent; no numerical input added

Resource	Inspected material	Decision and limit
Schultz 2026, DOI 10.1039/d5sc09493c; Zenodo 17151831	Publication and repository metadata	Fixed-cell chemistry candidate; archive not reanalysed
Feurerer 2021, DOI 10.1073/pnas.2113694118	Publication	Fixed-cell state precedent; no longitudinal input admitted
IJMS 2023, DOI 10.3390/ijms24010824	Publication and availability statement	Fixed samples and data on request; not reanalysed
Okumura 2026, DOI 10.1364/BOE.578769	Article and supplement	Murine model; not admitted as human evidence
Illinois resource, DOI 10.13012/B2IDB-5669414_V1	Metadata; repository route unsuccessful	Unresolved lead; access failure does not prove absence

The interior-crop and repeated-2950 analyses were implemented on 22 October after an editorial review of the initial package. Their selection used knowledge of earlier results and source descriptions. They are exploratory additions, not retrospective preregistration. The count of large video transitions was obtained from the largest ratio between sorted differences, not set to five from the source caption. The crop itself was chosen during inspection and remains a potential source of selection dependence.

The public supplement preserves the contemporaneous search-route record and a compact selection-history note. Later clarifications are identified as such; neither a missing complete query log nor a retrospective summary is presented as an exhaustive search history. No frequency of metadata limitations in the field can be estimated from this selection.

S3 Inputs and provenance

metadata/RAW_MANIFEST.json fixes five inputs by original URL, byte size and SHA-256. Their total size is 325,177,832 bytes. The four videos are included in Supplementary File S2; the 320,881,706-byte TIFF is recovered from its original repository. The downloader verifies existing files and rejects mismatches without replacing them. A historical redirect address is not a substitute for a new successful download.

Table S2. Decoded representations. Dimensions are height $\times$ width; playback rates describe containers, not biological acquisition cadence.

File	Frames or planes	Dimensions	Playback fps
canton_movie_2.mov	39 frames	405 $\times$ 319	7.0588
canton_movie_4.avi	40 frames	288 $\times$ 470	7.0000
canton_movie_6.avi	20 frames	512 $\times$ 512	7.0000
plos_s1_movie.bin	45 frames	480 $\times$ 640	30.3030
srs_Macrophages_5hCholOA_SRS.tiff	153 spectral planes	1024 $\times$ 1024	Not applicable

The .bin extension preserves the download name of a QuickTime movie. Display RGB equality was exact for all three Canton renderings; the PLOS overlay had unequal display values. Equal channels describe greyscale replication in those exports, not the original instrument's complete measurement channels. All complete decoded frames were byte-distinct within each movie, while the PLOS interior crops showed exact repetition.

Video source materials retain CC BY 4.0 and the SRS archive declares CC0. `metadata/SOURCE_LICENCES.json` records source citations, licence links and adaptations. Crops, montages and numerical summaries are attributed to the underlying datasets. Their creators are not authors of, or presumed to endorse, this secondary analysis.

S4 Time and evidence states

Five temporal properties were separated: playback time; acquisition cadence stated by the source; the mapping from acquisitions to rendered frames; acquisition origin; and statistical dependence between observations. The last is not resolved by a timestamp. Multiple images of the same field do not become independent biological replicates because their pixels differ.

Canton movie 2 has 39 rendered frames but a description of 15 minutes with 30-second acquisition intervals. A one-to-one consecutive-acquisition mapping is unresolved. Movies 4 and 6 have 40 and 20 frames and reported cadence of 15 seconds in nominal 10- and 5-minute windows. If every rendered frame corresponds to one consecutive acquisition, their first-to-last spans would be 585 and 285 seconds. Endpoint conventions can reconcile such spans with nominal windows; count compatibility is not proof of the mapping. An unknown absolute start does not itself invalidate relative cadence.

PLOS S1 is described as six images at five-minute intervals in a nominal 30-minute window. Six consecutive acquisitions at that cadence would span five intervals, or 25 minutes. The six recovered interior images are compatible with the reported image count but do not establish the first acquisition, endpoint convention or complete acquisition-to-display mapping. The overlaid clock is part of the export and was not accepted as an independent acquisition record.

Table S3. Definitions used for manual audit adjudication. These are task-specific evidence states, not a score or quality ranking.

State	Operational meaning	Example
Established	Supported directly by inspected data or sufficiently explicit source documentation	Exact crop repetition; documented spectral axis
Conditional	Derivable only under a stated, unverified assumption	Relative time using cadence and consecutive-acquisition mapping
Not established	Not annotated, not verified, or not sufficiently supported in this audit	An independent reference of cell correspondences
Absent in the representation	The required dimension is not encoded in the inspected representation	A live time sequence in the fixed SRS field

`metadata/audit_adjudications.csv` records the requirement, evidence, state, scope and attribution of each manual judgment. The checks were prepared with AI assistance; they do not assert an independently completed human annotation exercise. The author retains responsibility for the submitted interpretation. Boolean fields in the older video output named `known_independent_identity_labels` and `known_donor_ids_in_file` mean not established in the inspected materials, not proven absent from all records held by the original investigators. The explicit adjudication table governs that interpretation.

S5 Repeated interior content in PLOS S1

For consecutive complete frames, mean absolute difference is the sum of absolute differences across all $480 \times 640 \times 3$ stored values divided by their count. The interior calculation uses $384 \times 512 \times 3$ values from $x=64:576$ and $y=48:432$, with zero-based upper-exclusive bounds. No denoising, registration, intensity normalisation or interpolation was applied. Equality means identical decoded array values.

Table S4. Exact runs of repeated crop content. Run labels are bookkeeping identifiers for image arrays, not cell identities or biological states.

Run	Inclusive decoded indices	Frames
1	0–4	5
2	5–13	9
3	14–21	8
4	22–30	9
5	31–39	9
6	40–44	5

The 44 crop comparisons comprised 39 exact equalities and five changes. Each run was also checked against its first crop. The largest ratio between adjacent positive sorted full-frame differences was 208.38626553586533, with exact separation endpoints 0.05002495659722225 and 10.424513888888889. The next-largest ratio was 1.3401661070860575. A threshold strictly inside that open interval selects the same five boundaries. The geometric mean of the endpoints, 0.7221397751396934, was used for illustration only. Full-precision machine-readable outputs govern threshold comparisons; rounded endpoints must not be treated as exact bounds.

The crop and the separation rule use overlapping information from the same movie. Their agreement is not independent biological evidence. The result does not establish repetition of all peripheral pixels, original acquisition events, physical timing, incorporation completion or cell correspondence. Supplementary Figure S1 shows every complete rendered frame with the crop boundary for inspection.

S6 Spectral reconstruction and the repeated coordinate

The source README supplies 153 Raman-shift entries and background values. There are three acquisition blocks, each containing 51 planes: 1625–1775, 2800–2950 and 2950–3100 cm^{-1} at 3 cm^{-1} increments. Planes 101 and 102 therefore share nominal 2950 cm^{-1} . They were not dropped, merged or assigned a common calibrated amplitude. The fixed sample, spectral dimension and figure-panel directory are distinct metadata facts.

Table S5. Repeated-setting comparison in one fixed field. Means use all 1,048,576 pixels, including background. Plane indices are zero-based. Units are arbitrary stored-signal units.

Quantity	Block 2 plane 101	Block 3 plane 102
Nominal Raman shift in cm^{-1}	2950	2950
Raw field mean	211.38975143432617	244.16172313690186
Source background	154.929756756757	167.750202702703
Signed mean minus background	56.45999467756917	76.41152043419885

The raw difference is 32.771971702575684 units, or 15.50310338141306% of the block-2 raw mean. After background subtraction, the difference is 19.951525756629678 units, or 35.33745596429566% of the corresponding block-2 value. Both use $100 \times (\text{block 3} - \text{block 2}) / \text{block 2}$. The changed denominator helps explain why the percentages differ; neither is a population effect or concentration ratio. The background was supplied by the source, not estimated through a new calibration.

The main script decodes all 153 pages separately and verifies field means through integer sums. An alternative verification reads the two focal planes with Pillow and computes means from 65,536-bin histograms with integer products and sums. Those two means agree exactly. The original full-array SRS analysis additionally compared separately decoded pages and reproduced standard deviations by second moments, with maximum absolute difference 2.56×10^{-13} stored-signal units. These calculations test reading and arithmetic, not the cause of the between-block discrepancy.

S7 Supplementary spatial and region illustrations

Maps at 2857, 2860 and 2875 cm^{-1} come from the same acquisition block. A shared display range of 0–1674 units derives from pooled 0.5–99.5 percentiles. Clipping affects display only; calculations retain original values. Coordinates remain pixels because physical pixel calibration was not established from the inspected TIFF tags. Neither a lack of maximum stored values nor the maps establish absence of detector saturation.

Region A is the 32×32 tile with greatest mean across the three display planes. Region B, excluding A, is closest to the median tile mean; ties follow row-major order. The candidate grid contains 1,024 tiles. A occupies $x=672-703$, $y=736-767$; B occupies $x=448-479$, $y=800-831$, with inclusive bounds. B can include predominantly background. Its selection does not make it a median cell. The raw 2860 cm^{-1} means, 1616.61 and 193.33, reflect a contrast partly ensured by selection.

The spatial maps and block-specific profiles are retained to make the extraction inspectable, not as independent evidence of cellular classes, organelle identity or lipid composition. No molecular species, lipid origin or phenotype was inferred from these regions. A separate exploratory grid-sensitivity exercise in the local working history is not a result of this submission and is not needed to reproduce the reported comparisons.

S8 Numerical output guide

Table S6. Main machine-readable outputs. Paths are relative to the extracted Supplementary File S2.

Output	Records	Interpretation
results/video_audit.csv	Four movies	Container and decoded-array properties; manual flags require adjudication table
results/srs_spectral_summary.csv	153 planes	Mean, spatial spread, quantiles, extremes and documented background
results/srs_two_regions_spectra.csv	306 region-plane rows	Two selected tiles across three blocks; not cellular replicates
results/srs_all_tile_selection_scores.csv	1,024 tiles	Complete candidate grid used for A/B selection
results/representation_findings/plos_transition_differences.csv	44 pairs	Full-frame and crop differences; exact equality
results/representation_findings/plos_visual_runs.csv	Six runs	Repeated-content intervals
results/representation_findings/plos_sorted_full_frame_differences.csv	44 differences	Ordered differences and adjacent ratios
results/representation_findings/srs_field_means_by_block.csv	153 planes	Reconstructed axes, means and backgrounds
results/representation_findings/srs_repeated_2950.csv	Two planes	Focal repeated-setting comparison

JSON summaries repeat selected results and calculation limits; they are not additional datasets.

metadata/srs_spectral_axis_and_background.csv exposes the source axis and background. Video differences are in stored 8-bit units; SRS amplitudes are in arbitrary stored-signal units. Quantiles and standard deviations

describe spatial variation, not confidence intervals across donors. Dates, logs, paths and provenance receipts can change on rerun without changing numerical results.

S9 Reproduction and verification

Extract Supplementary File S2 to a writable directory. Create and activate a Python 3.12 virtual environment, install the pinned requirements and execute the commands below from the extracted root. Scripts also resolve paths from their own locations. No authentication or remote compute service is required.

```
python -m pip install -r requirements.txt
python code/fetch_verified_raw.py
python code/inspect_videos.py
python code/analyse_video_audit.py
python code/srs_analysis.py
python code/analyse_representation_findings.py
python code/verify_representation_findings.py
```

Use one downloader per destination. Existing originals must match the manifest; mismatches are rejected.

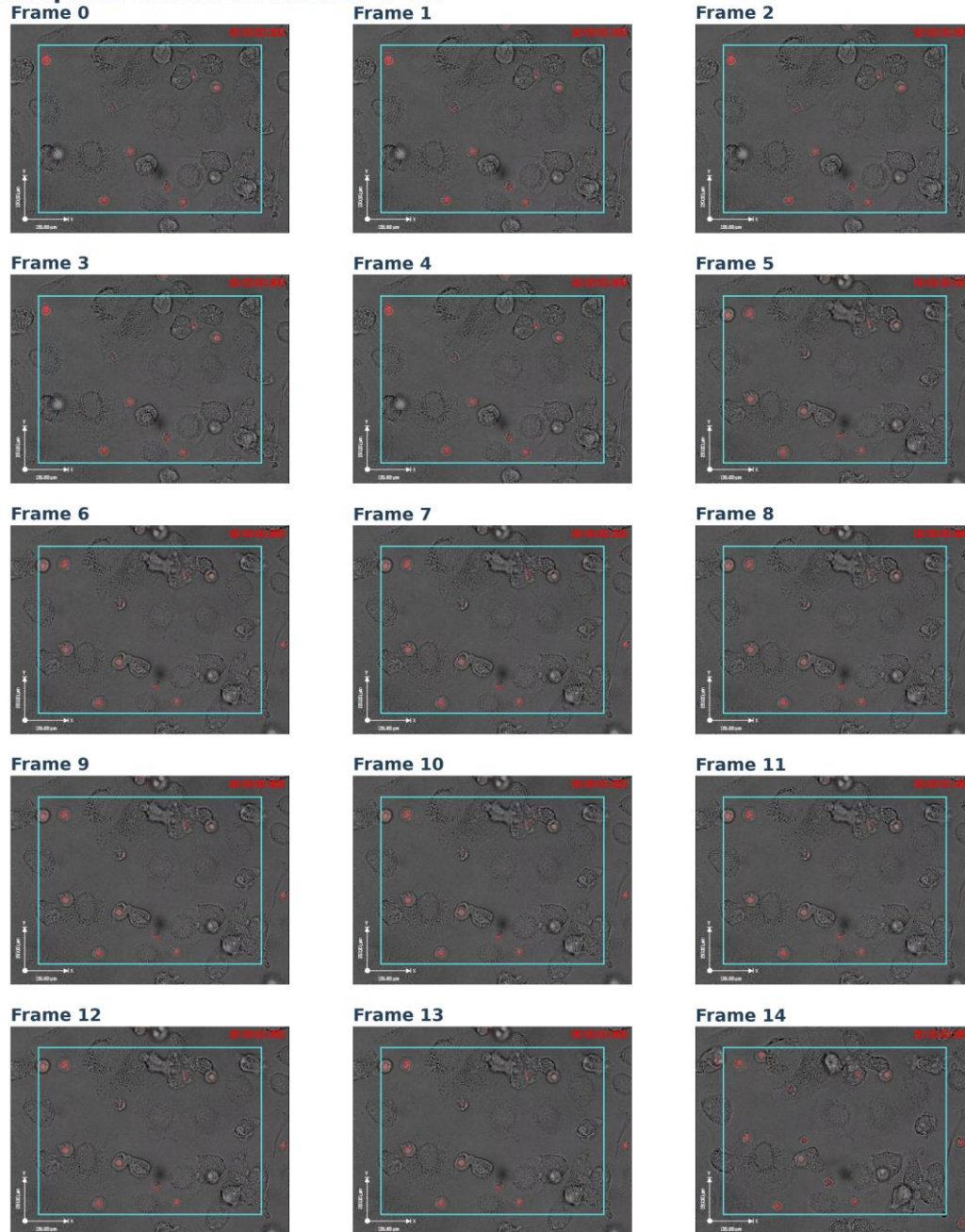
`python code/fetch_verified_raw.py --verify-only` checks originals without network access. The SRS analysis needs memory beyond the 321 MB TIFF for decoded arrays and temporary calculations. Decoded arrays are regenerated in `data/derived/`; reference outputs are in `results/` and `figures/`. Regeneration overwrites derived outputs in the extracted working copy, so retain the distributed ZIP for comparison.

The revised video audit rejects nonzero FFmpeg return codes, missing completion or count information, and disagreement with the primary decoded count before writing accepted tables or figures. A simulated failure in an isolated test confirmed that no new audit tables or figures were produced. The normal counts were 39, 40, 20 and 45. Both reading routes use FFmpeg components; a second interface is not an independent codec implementation.

The main repetition analysis and an alternative FFmpeg RGB reconstruction agree on six crop arrays, 39 equalities and five boundaries. Pillow and integer-histogram arithmetic independently reproduce the focal SRS means relative to the tiff file route. The verification receipts identify scripts, input hashes, library versions and scope. A package rerun compares scientific tables and numerical JSON separately from timestamped provenance, and PNG pixels separately from PDF metadata. The current record is `verification/REPRODUCTION_CHECK_MAJOR_REVISION.json`; the previous Windows check remains explicitly historical.

The package verification concerns the stated local software environment and the same biological inputs. It does not establish universal cross-platform pixel identity. No computation in this supplement constitutes a new biological replicate.

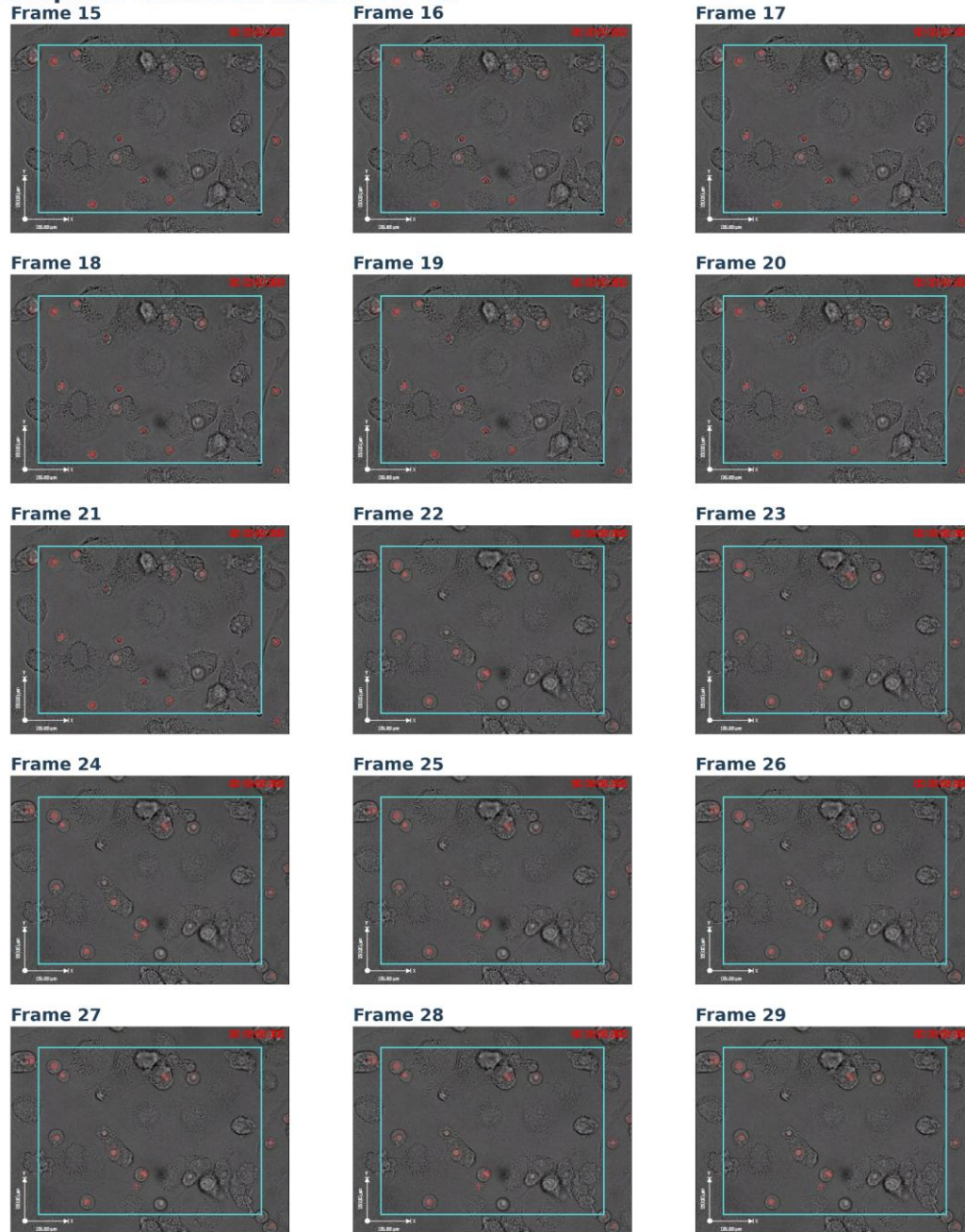
Complete rendered frames 0-14



Cyan rectangle: analysed crop $x = 64:576$, $y = 48:432$; upper bounds excluded.
Original clock and scale overlays retained. Frame index does not certify acquisition time.
No identity or trajectory annotations. Source: Santulli-Marotto et al., PLOS ONE (2015),
S1 Movie, CC BY 4.0. Decoded, assembled and overlaid with the crop boundary.

Supplementary Figure S1A. Complete PLOS rendering and crop boundaries. Parts A–C show all 45 decoded frames in order, preserving the original overlays. Rectangles mark the exploratory crop $x=64:576$, $y=48:432$; zero-based indices are printed outside the source images. The added boundary is a crop annotation, not a cell mask or trajectory. Source: Santulli-Marotto et al., PLOS ONE 2015, DOI 10.1371/journal.pone.0145078, S1 Movie, CC BY 4.0. The complete original movie is included in S2.

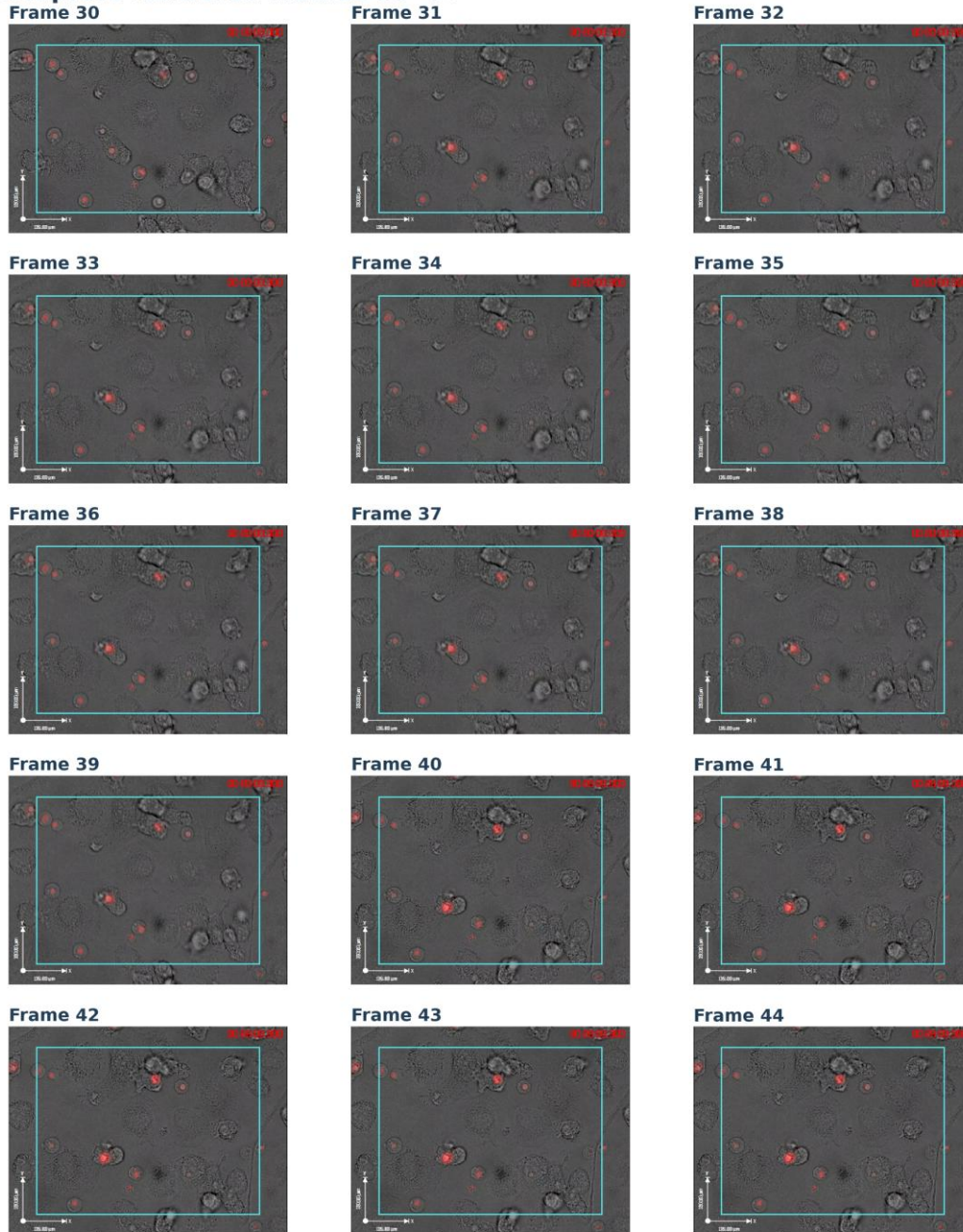
Complete rendered frames 15-29



Cyan rectangle: analysed crop x = 64:576, y = 48:432; upper bounds excluded.
Original clock and scale overlays retained. Frame index does not certify acquisition time.
No identity or trajectory annotations. Source: Santulli-Marotto et al., PLOS ONE (2015), S1 Movie, CC BY 4.0. Decoded, assembled and overlaid with the crop boundary.

Supplementary Figure S1B. Complete PLOS rendering and crop boundaries. Parts A–C show all 45 decoded frames in order, preserving the original overlays. Rectangles mark the exploratory crop x=64:576, y=48:432; zero-based indices are printed outside the source images. The added boundary is a crop annotation, not a cell mask or trajectory. Source: Santulli-Marotto et al., PLOS ONE 2015, DOI 10.1371/journal.pone.0145078, S1 Movie, CC BY 4.0. The complete original movie is included in S2.

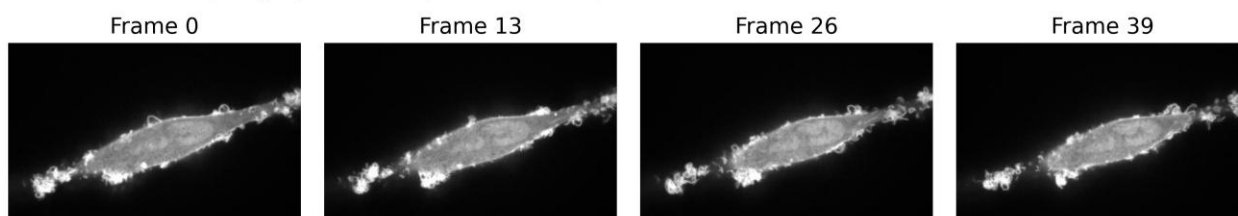
Complete rendered frames 30-44



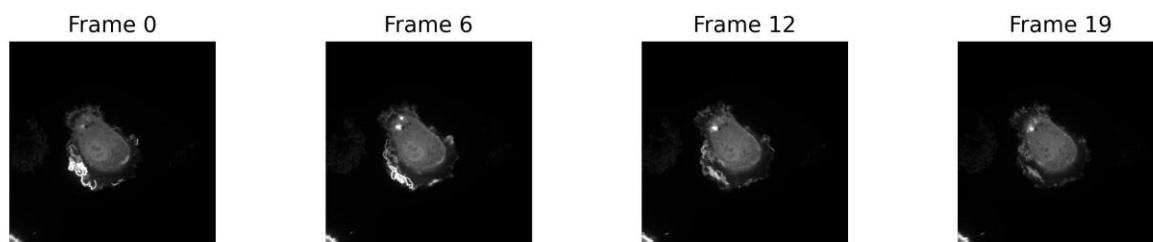
Cyan rectangle: analysed crop $x = 64:576$, $y = 48:432$; upper bounds excluded.
 Original clock and scale overlays retained. Frame index does not certify acquisition time.
 No identity or trajectory annotations. Source: Santulli-Marotto et al., PLOS ONE (2015),
 S1 Movie, CC BY 4.0. Decoded, assembled and overlaid with the crop boundary.

Supplementary Figure S1C. Complete PLOS rendering and crop boundaries. Parts A–C show all 45 decoded frames in order, preserving the original overlays. Rectangles mark the exploratory crop $x=64:576$, $y=48:432$; zero-based indices are printed outside the source images. The added boundary is a crop annotation, not a cell mask or trajectory. Source: Santulli-Marotto et al., PLOS ONE 2015, DOI 10.1371/journal.pone.0145078, S1 Movie, CC BY 4.0. The complete original movie is included in S2.

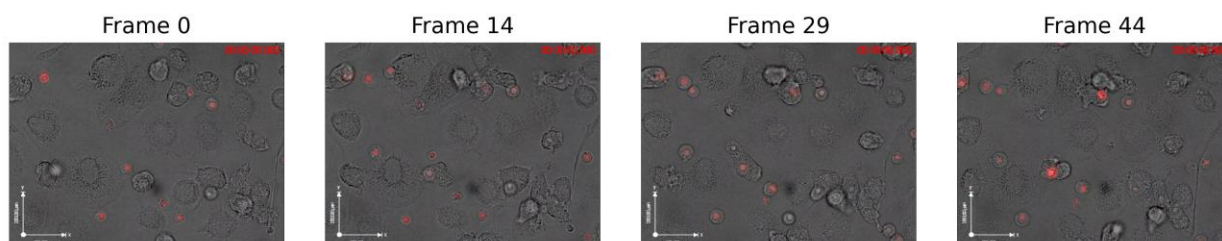
A Human macrophage | confocal | Canton 2016, movie 4



B Human macrophage | confocal | Canton 2016, movie 6



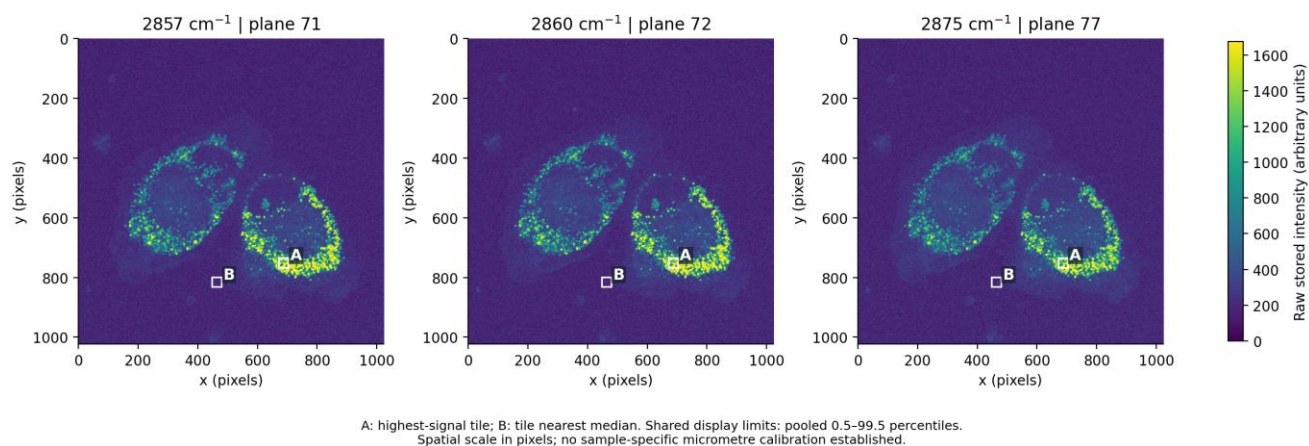
C Human efferocytosis | Santulli-Marotto 2015, S1



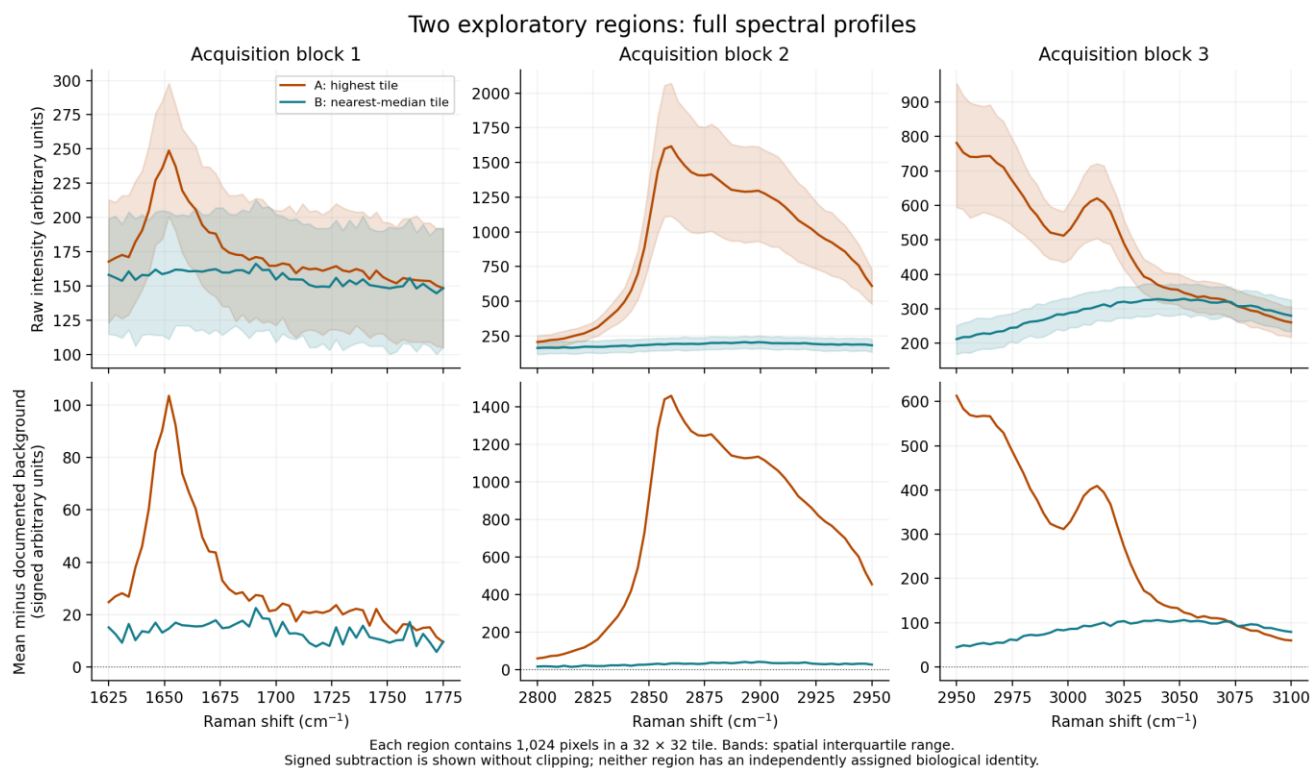
Published rendered videos; fields selected by the original authors.
Frame indices do not establish acquisition times. No raw spectra or
independent identity annotations are provided by these exported files.

Supplementary Figure S2. Illustrative live human video frames. Selected decoded indices from Canton movies 4 and 6 and Santulli-Marotto S1, assembled without new cell correspondence or event annotations. Original fluorescent/colour display overlays are preserved. Canton et al., DOI 10.1038/ncomms11284, and Santulli-Marotto et al., DOI 10.1371/journal.pone.0145078, CC BY 4.0. These are renderings, not recovered independent spectral channels. Their complete source files and counts are in S2.

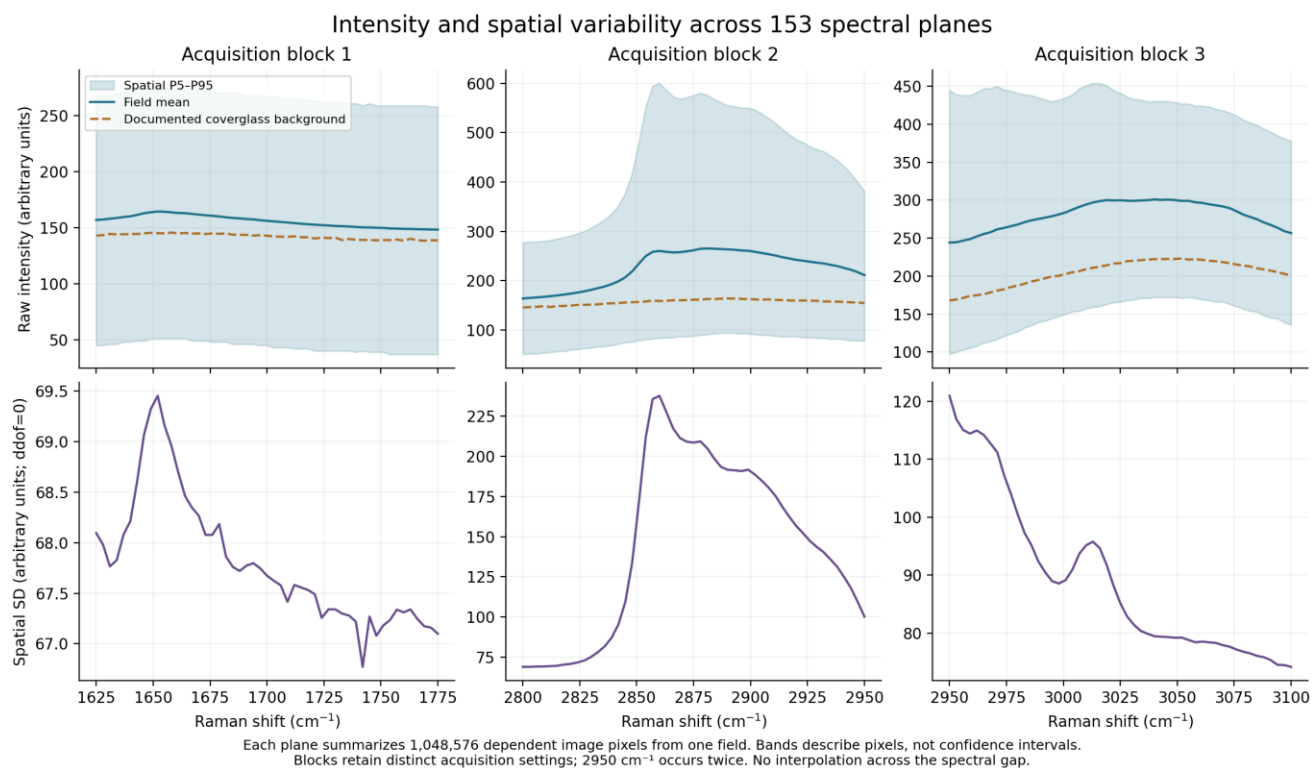
Spatial spectral contrast in one fixed human macrophage field



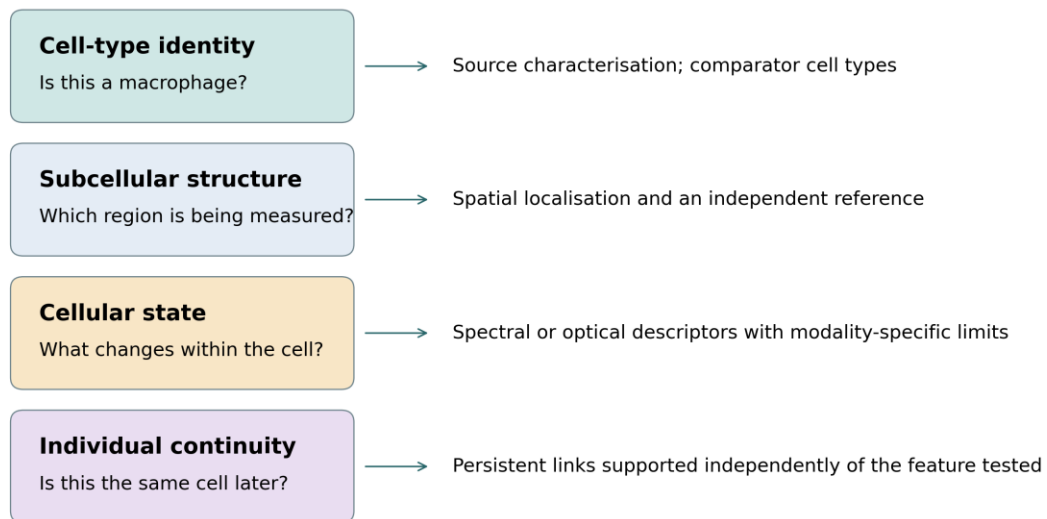
Supplementary Figure S3. Spatial maps and intensity-selected tiles. Fixed human SRS field at 2857, 2860 and 2875 cm^{-1} , displayed with the shared range described in S7. Coordinates are pixels, not physical distance. A maximises and B approximates the median tile score; neither is an identified cell. Source: S-BIAD1584, DOI 10.6019/S-BIAD1584, CC0, associated with panel 3D of Szkalitsity et al. 2025.



Supplementary Figure S4. Profiles of intensity-selected regions. A/B profiles retain the three acquisition blocks and both entries at 2950 cm⁻¹. Selection-induced amplitude differences are descriptive and do not test biological classes. B can contain background. Source: S-BIAD1584, CC0.



Supplementary Figure S5. Complete field spectral summaries. All 153 planes appear by block. Spatial spread refers to pixels within one field, not biological-replicate uncertainty. Raw field means and documented coverglass background are separate; signed subtractions are in the output tables. No full molecular decomposition or concentration calibration was reproduced. Source: S-BIAD1584, CC0.



The same cell may change its spectrum; different cells may share similar spectra.

Conceptual framework. The arrows mean "requires supporting evidence", not a biological causal chain.

Supplementary Figure S6. Four questions of identification. Retained conceptual diagram distinguishing cell-type attribution, subcellular location, measured state and continuity of an individual. Arrows represent evidence requirements, not a biological causal chain. The schematic is explanatory rather than empirical.