Supplementary Figures

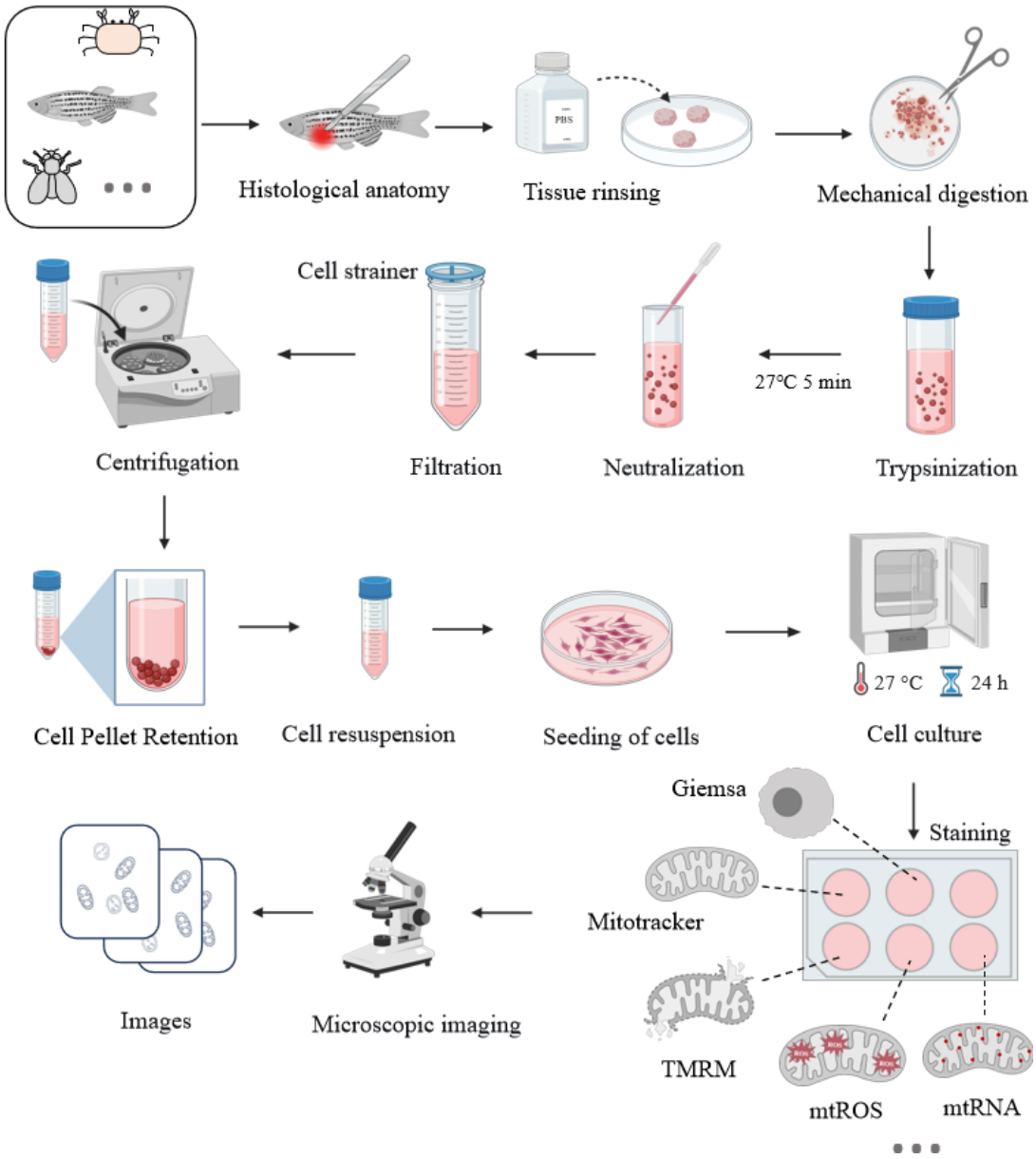


Fig. S1 – Workflow for primary cell isolation and mitochondrial fluorescent labeling. Following anesthesia and surface disinfection of the animals, target tissues were aseptically isolated. The tissue fragments were rinsed with antibiotic-containing 1×PBS, minced, and subjected to low-speed centrifugation at 50 × g for 5 min. The resulting pellet was digested with 0.25% trypsin until a flocculent consistency was observed. Digestion was terminated by adding complete culture medium supplemented with 20% serum. The cell suspension was filtered through a 100 μm cell strainer, centrifuged, and resuspended, followed by plating at an appropriate density for culture. Cells were stained with MitoTracker or alternative dyes to visualize mitochondrial networks and with Giemsa staining for cellular morphology observation, followed by microscopic imaging and data analysis.

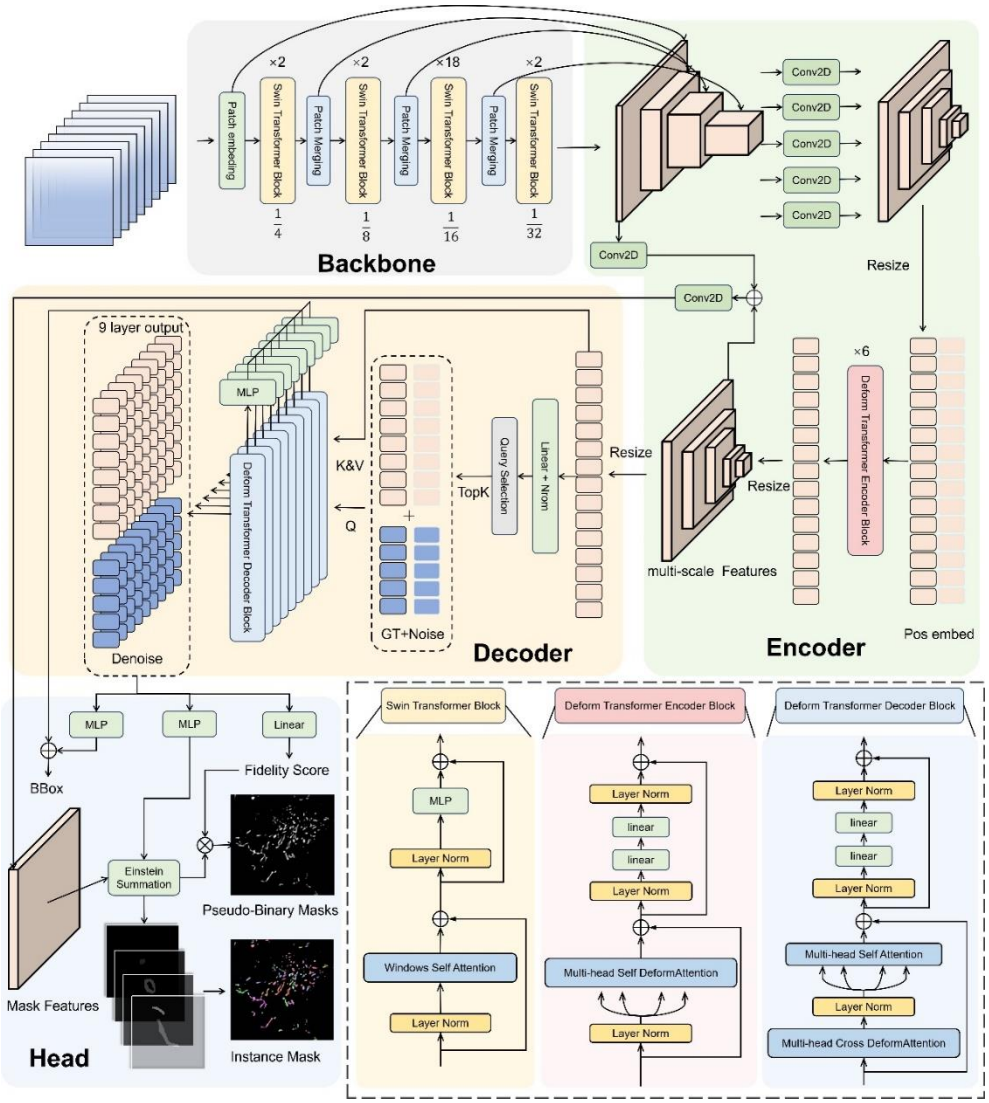


Fig. S2 – Architecture of PanoMitoSeg. The network comprises four main components: Backbone, Decoder, Encoder, and Head. **Backbone:** Four Swin Transformer modules serve as the feature extraction backbone. Each module is followed by patch merging, which performs 2× down-sampling to generate multi-scale feature representations. **Encoder:** A convolutional module normalizes the channel dimensions of the multi-scale features. The channel-aligned features are then flattened into tokens and fed into a 6-layer multi-head deformable transformer module to further enhance the multi-scale feature representation. **Decoder:** Multi-scale features are flattened into tokens, from which a set of Top-K query vectors is selected. A mixed query strategy is applied to these queries and all tokens. A denoising training strategy improves robustness. The Decoder iteratively refines the output through a 9-layer multi-head deformable transformer module. Each output of the Decoder is passed to the Head and directly optimized by the loss function. **Head:** The Head consists of three output branches: a box head, a fidelity-score head, and a mask head. The mask head produces weight vectors that are used to weight the mask features, which is the sum of the highest-resolution features from the Backbone and the Encoder, to generate instance segmentation masks, which are then combined into an instances mask. Each instance mask is further weighted by its corresponding fidelity score to produce the final pseudo-semantic map for duplicate suppression during training.

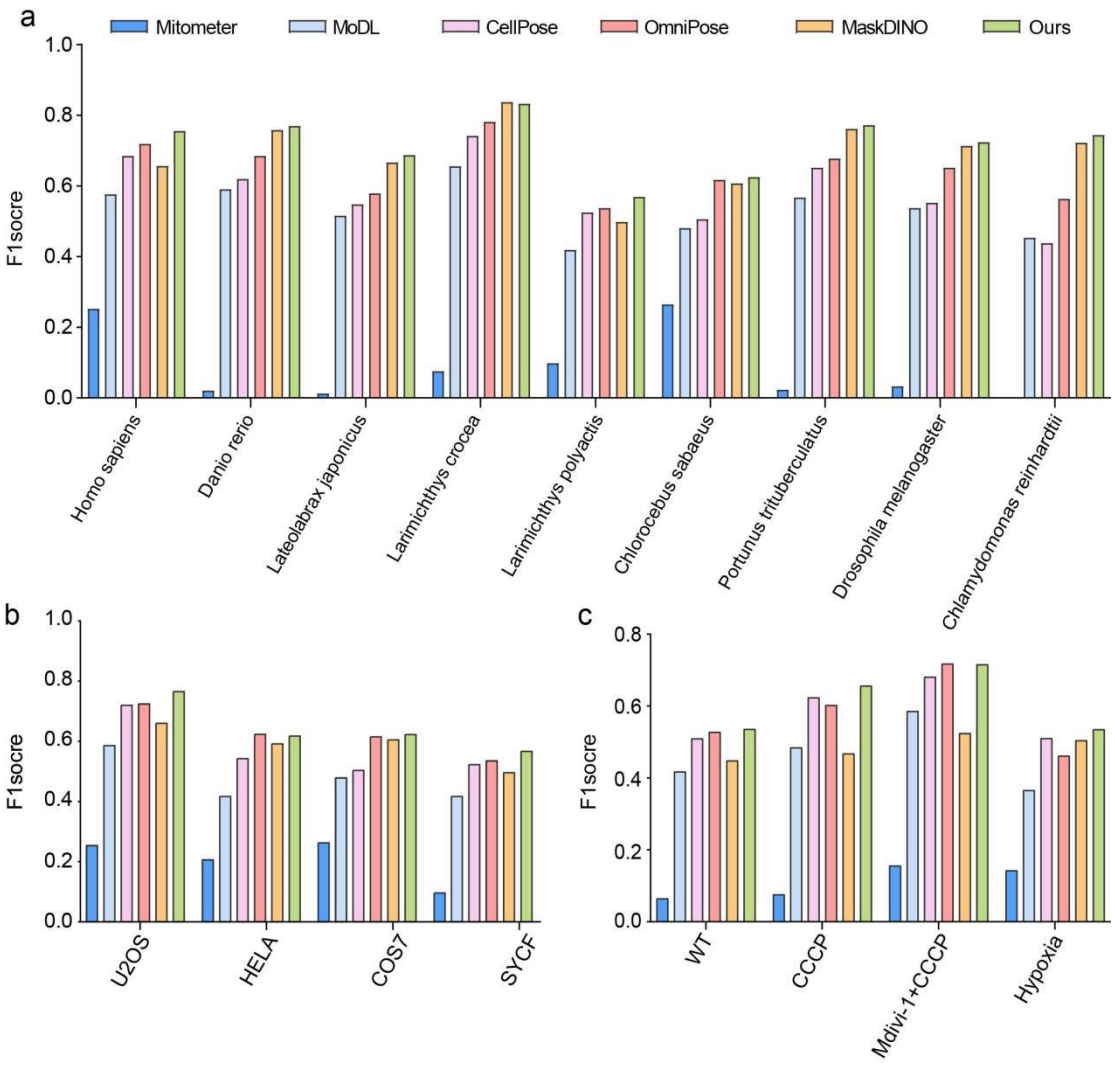


Fig. S3 – Performance of benchmarked methods across species, cell lines, and treatment conditions subsets. F1-score comparison of PanoMitoSeg and benchmarked methods, including MitoMeter, MoDL, CellPose-SAM, OmniPose, and MaskDINO, across **(a)** different species, **(b)** cell lines and **(c)** treatment subsets in PanoMitoAtlas. Each group corresponds to one biological subset, and each bar represents the F1-score of one method.

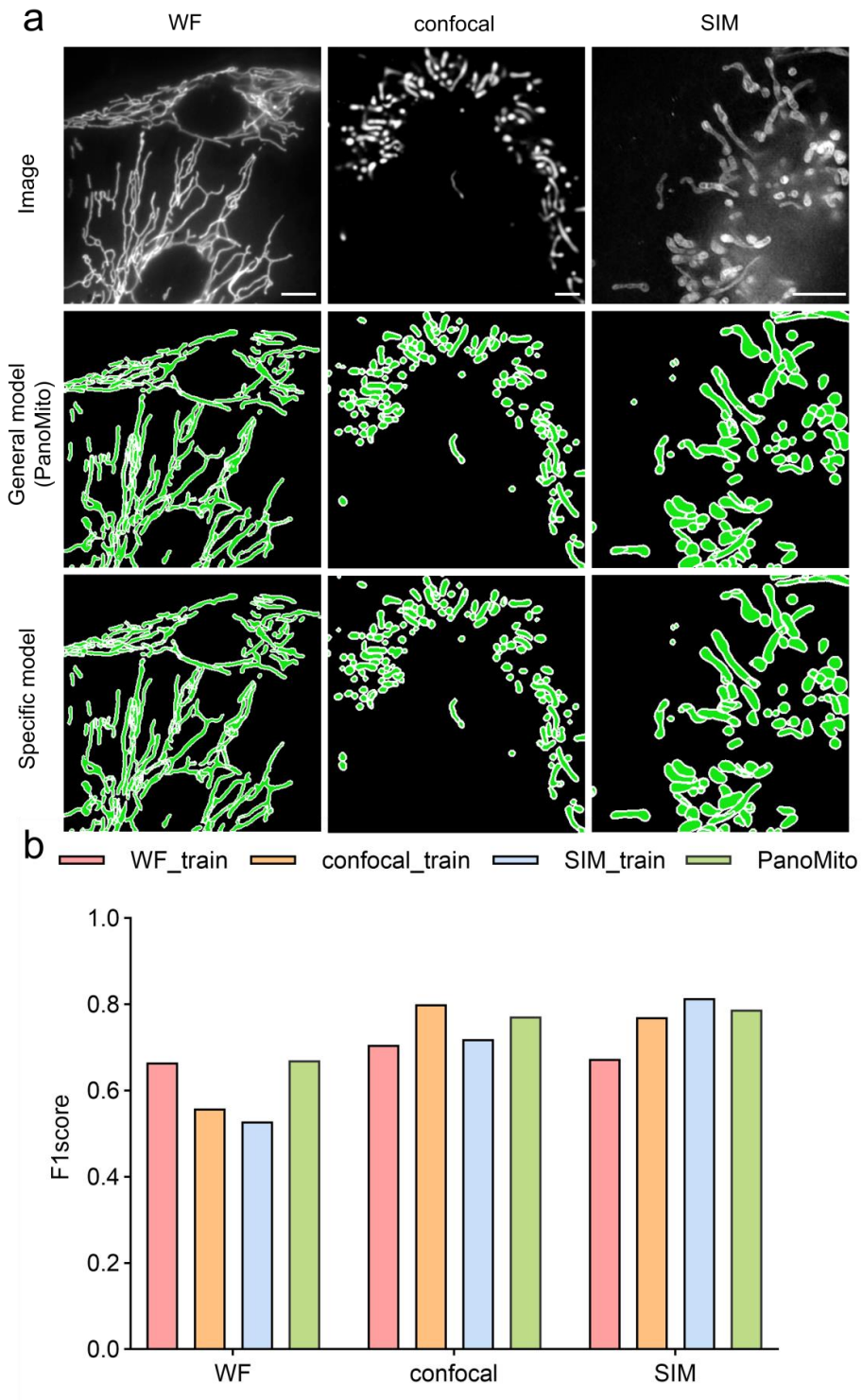


Fig. S4 – Modality-stratified validation of PanoMitoSeg. (a) Representative raw images and corresponding segmentation results from three imaging modalities, including wide-field fluorescence microscopy (WF), confocal microscopy, and SIM. The general model refers to PanoMito trained on the full multi-modality dataset, whereas the specific model refers to the PanoMito model trained only on data from the corresponding imaging modality. **(b)** Modality-stratified validation of PanoMitoSeg. Performance comparison of modality-specific models (WF_train, confocal_train, and SIM_train) and the general model (PanoMito) across WF, confocal, and SIM test data, reported as F1-score. Scale bar: 5 μm.

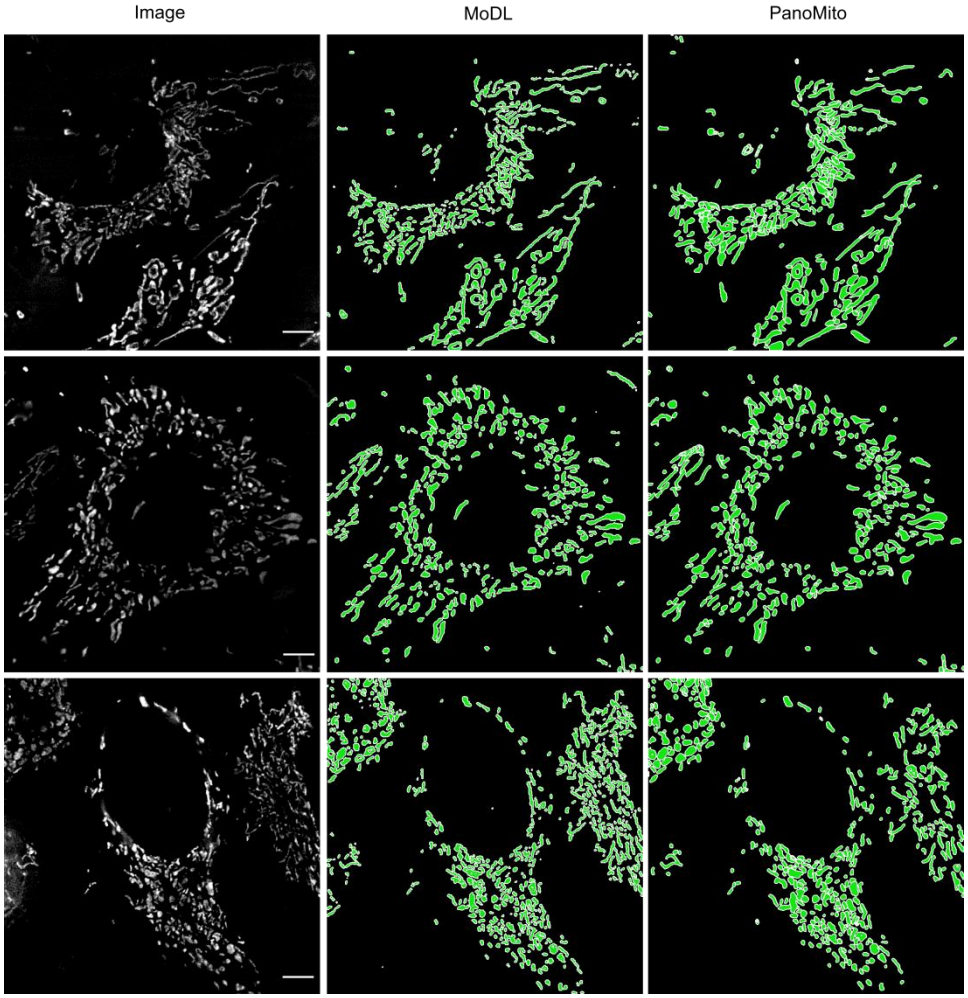


Fig. S5 – Representative examples from the densely packed mitochondrial subset. Representative images from the challenging subset characterized by densely packed and structurally complex mitochondria. The left column shows raw mitochondrial images, the middle column shows the corresponding MoDL segmentation results, and the right column shows the corresponding PanoMito segmentation results. Scale bar: 5 μm.

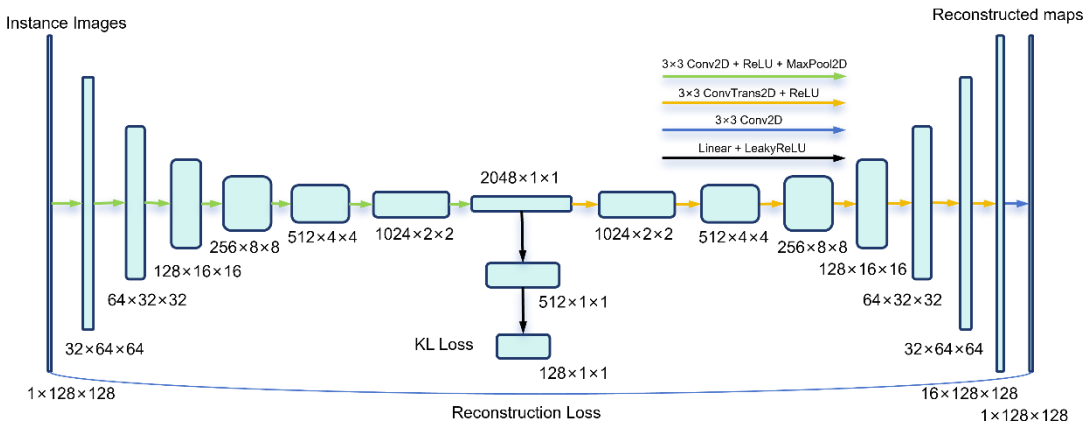


Fig. S6 – Architecture of PanoMitoCluster. The network consists of a CNN based encoder-decoder architecture with a self-supervised reconstruction objective. The bottleneck features are processed through two fully connected layers, with a Kullback–Leibler (KL)-divergence loss applied to the output distribution to enhance cluster separability. The final 128-dimensional feature vectors are used for k-means clustering of mitochondrial phenotypes.

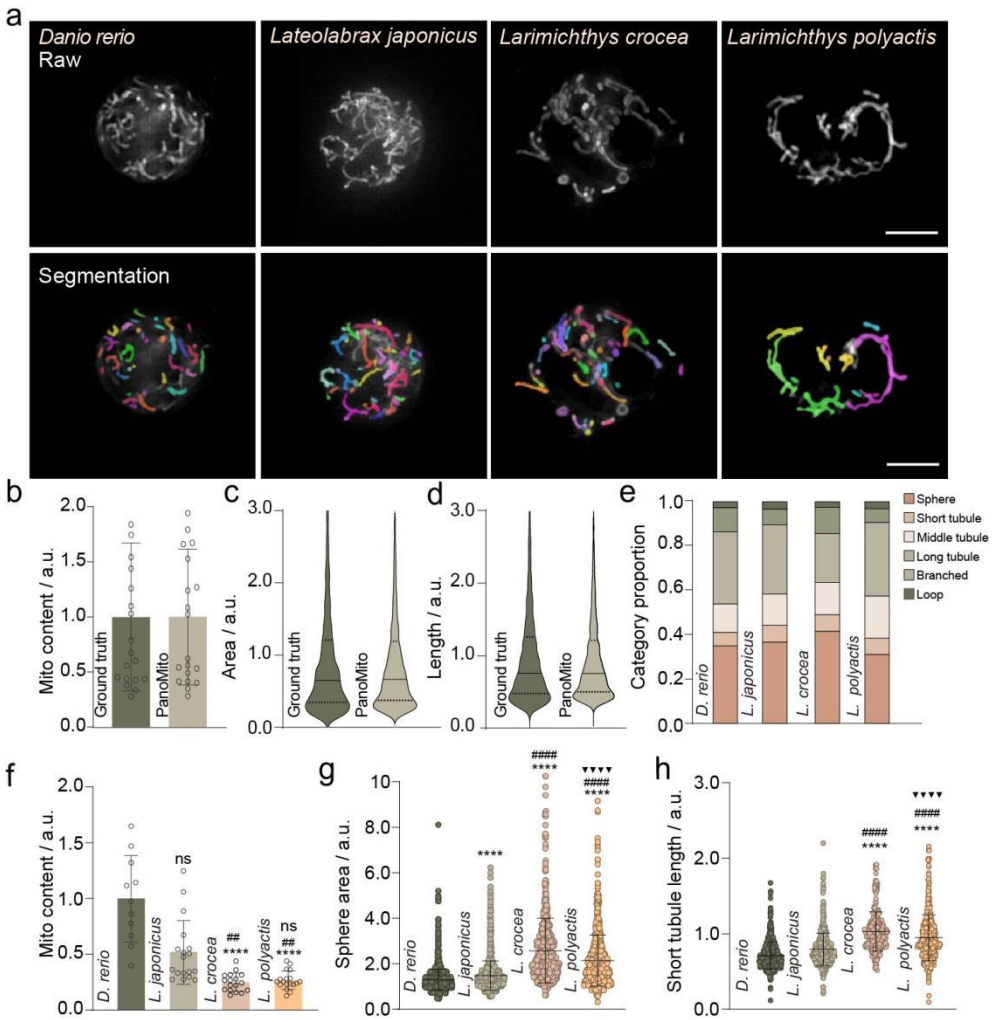


Fig. S7 – Intracellular mitochondrial heterogeneity across species. (a) Representative fluorescence and instance segmentation images of mitochondrial networks in oogonia from freshwater fish (*D. rerio* and *L. japonicus*) and marine fish (*L. crocea* and *L. polyactis*). Scale bars: 10 µm. **(b-d)** Accuracy assessment of PanoMitoNet versus manual annotations: mitochondrial count (b), area (c), and length (d). **(e–h)** Quantitative analysis of mitochondrial morphology in oogonia ($n \geq 20$ per species): (e) proportion of six morphological categories; (f) relative mitochondrial counts per cell; (g) relative sphere area; (h) relative short tubule length. Data: mean ± s.d. * vs first group; # vs second group; ▲ vs third group ($p < 0.05$).

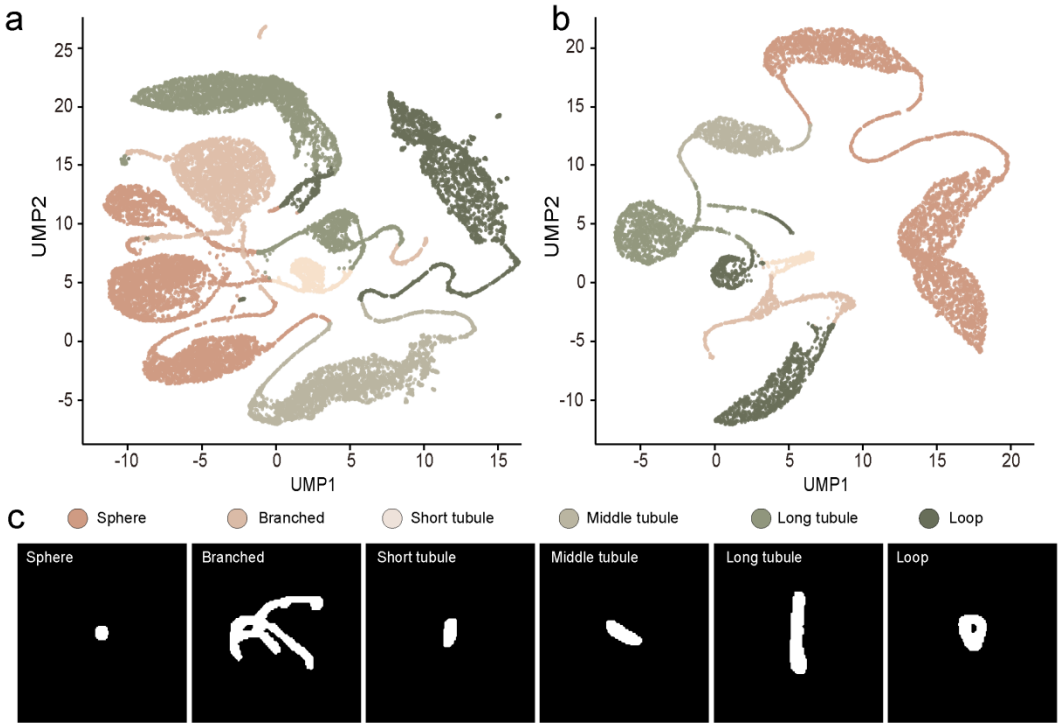


Fig. S8 – UMAP visualization of mitochondrial morphological heterogeneity. K-means clustering was performed on 128-dimensional feature vectors to initially identify 12 clusters; adjacent clusters were subsequently assigned using a deterministic hierarchical rule based on cluster-level morphological features to yield 6 final morphological classes for biological interpretation. Colors represent distinct clusters. **(a)** UMAP projection of mitochondria from oogonia across fish species. **(b)** UMAP projection of mitochondria from immune cells in the zebrafish kidney. **(c)** Representative instances illustrating the different morphologies of mitochondrial clusters.

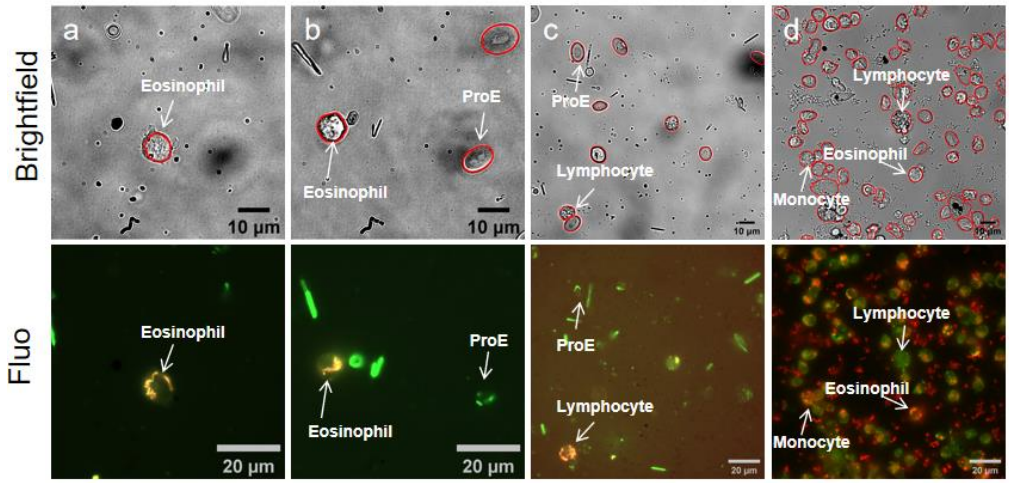


Fig. S9 – Brightfield and corresponding fluorescence images of immune cells in the zebrafish kidney. (a–c) Representative brightfield and mitochondrial membrane potential ($\Delta\Psi_m$) fluorescence images showing (a) a single cell, (b) two cells, and (c) multiple cells per field of view (1024×1024 for a, b; 2048×2048 for c). **(d)** Representative brightfield and *nd4* mRNA fluorescence images showing multiple cells per field of view (2048×2048). In all panels, extracellular red signals are debris/impurities, and red circles indicate individual cells segmented by the fine-tuned CellPose model.

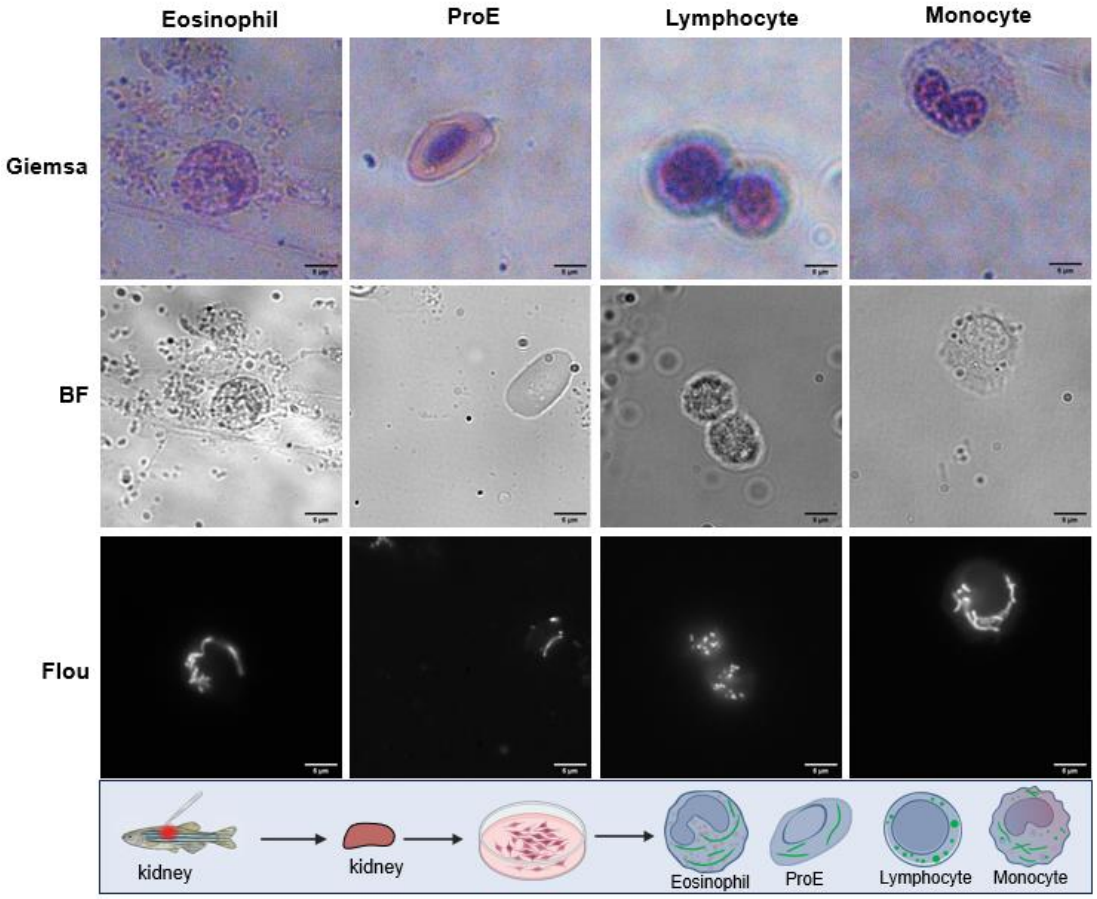


Fig. S10 – Identification of immune cell types in adult zebrafish kidney. Representative images of eosinophils, primitive erythrocytes (ProE), lymphocytes, and monocytes (left to right). For each cell type, four panels are shown (top to bottom): Giemsa staining, bright-field, MitoTracker Green fluorescence (mitochondria), and schematic illustration of cellular morphology. Immune cell types were classified based on established morphological criteria from Giemsa staining. Scale bar: 5 μm.

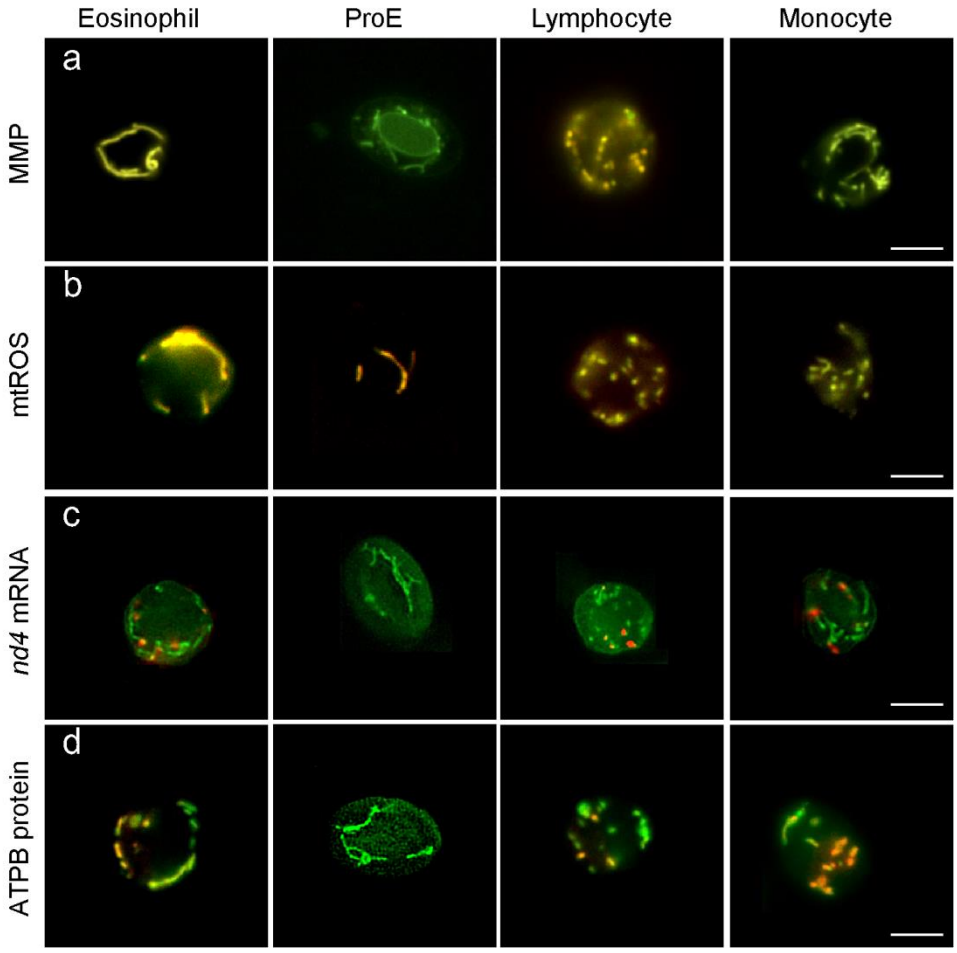


Fig. S11 – Raw fluorescent images of mitochondrial function and molecular expression across immune cell types in zebrafish kidney (corresponding to Fig. 3b–e). Representative images showing **(a)** mitochondrial membrane potential ($\Delta\Psi_m$), **(b)** mitochondrial ROS (mtROS), **(c)** *nd4* mRNA (red puncta), and **(d)** ATPB protein (orange puncta) in eosinophils, primitive erythrocytes, lymphocytes, and monocytes (left to right). Scale bar: 5 μm.

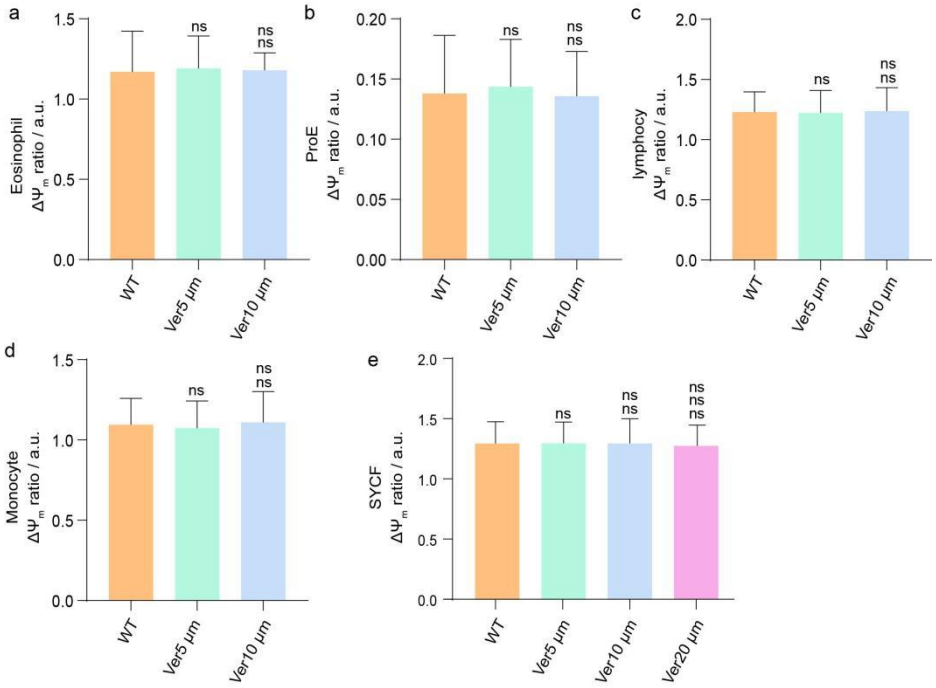


Fig. S12 – Dye-efflux control experiments using verapamil. (a–d) Relative $\Delta\Psi_m$ in zebrafish kidney immune cells after treatment with different concentrations of verapamil (0, 5, 10 µM): (a) eosinophils, (b) primitive erythrocytes (ProE), (c) lymphocytes, and (d) monocytes. **(e)** Relative $\Delta\Psi_m$ in SYCF cells after treatment with different concentrations of verapamil (0, 5, 10, 20 µM). Data are shown as mean ± s.d. No significant differences were observed between verapamil treated groups and the untreated control (WT) for any cell type (ns, $p > 0.05$).

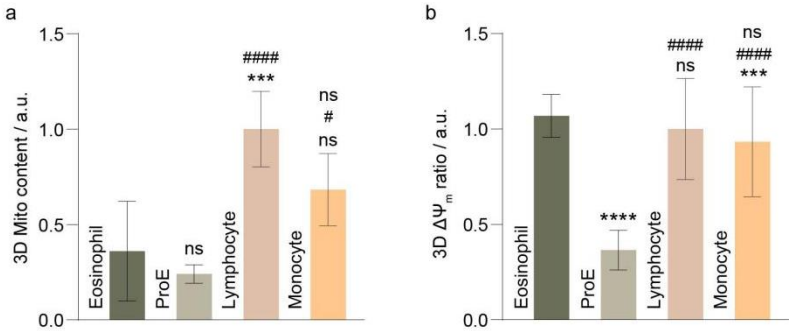


Fig. S13 – Quantitative 3D analysis of (a) mitochondrial count per cell and (b) mitochondrial membrane potential ($\Delta\Psi_m$) in immune cells. Data are shown as mean ± s.d. (n ⩾ 5 per cell type). Significance ($p < 0.05$) for pairwise comparisons is denoted as follows: * vs the first group; # vs the second group; ▲ vs the third group. The statistical comparisons based on 3D-annotated ground truth closely matched those obtained from 2D segmentation in Fig. 3f, h.

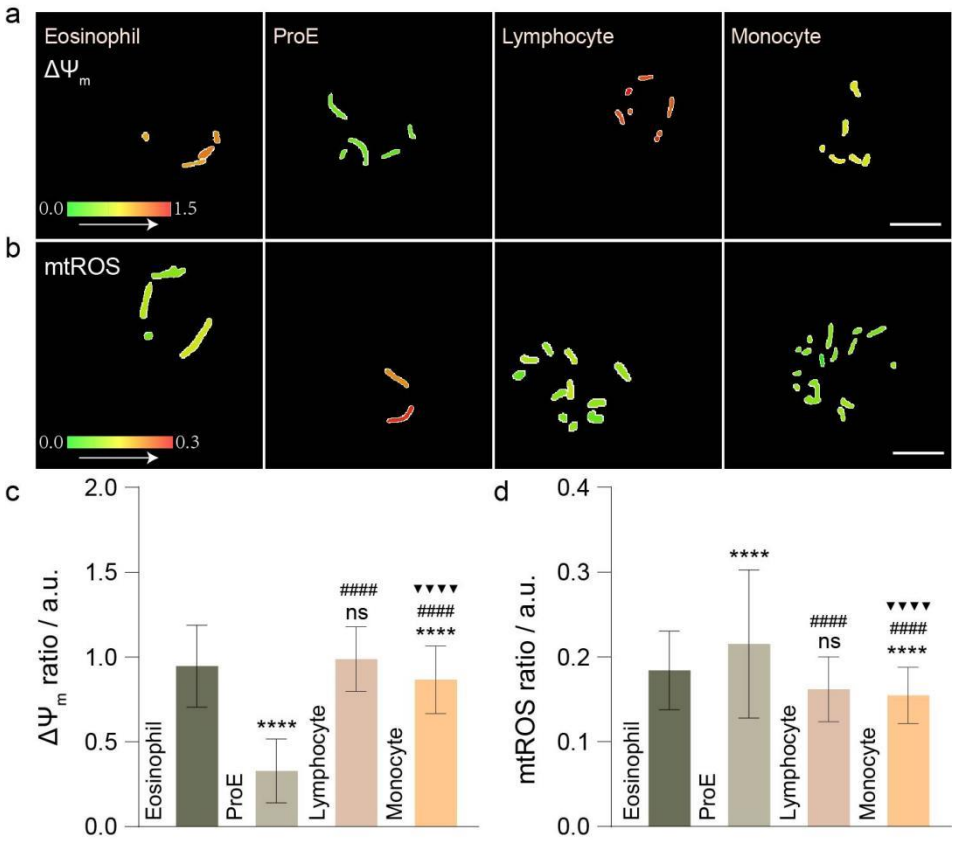


Fig. S14 – Intracellular mitochondrial heterogeneity across cell types in *L. japonicus* kidneys. **(a–b)** Representative structured images of mitochondrial function in *L. japonicus* kidney immune cells: (a) $\Delta\Psi_m$, (b) mtROS. Scale bar: 5 µm. **(c-d)** Quantitative function in immune cells: (c) $\Delta\Psi_m$, (d) mtROS. Data: mean ± s.d. * vs the first group; # vs the second group; ▲ vs the third group ($p < 0.05$).

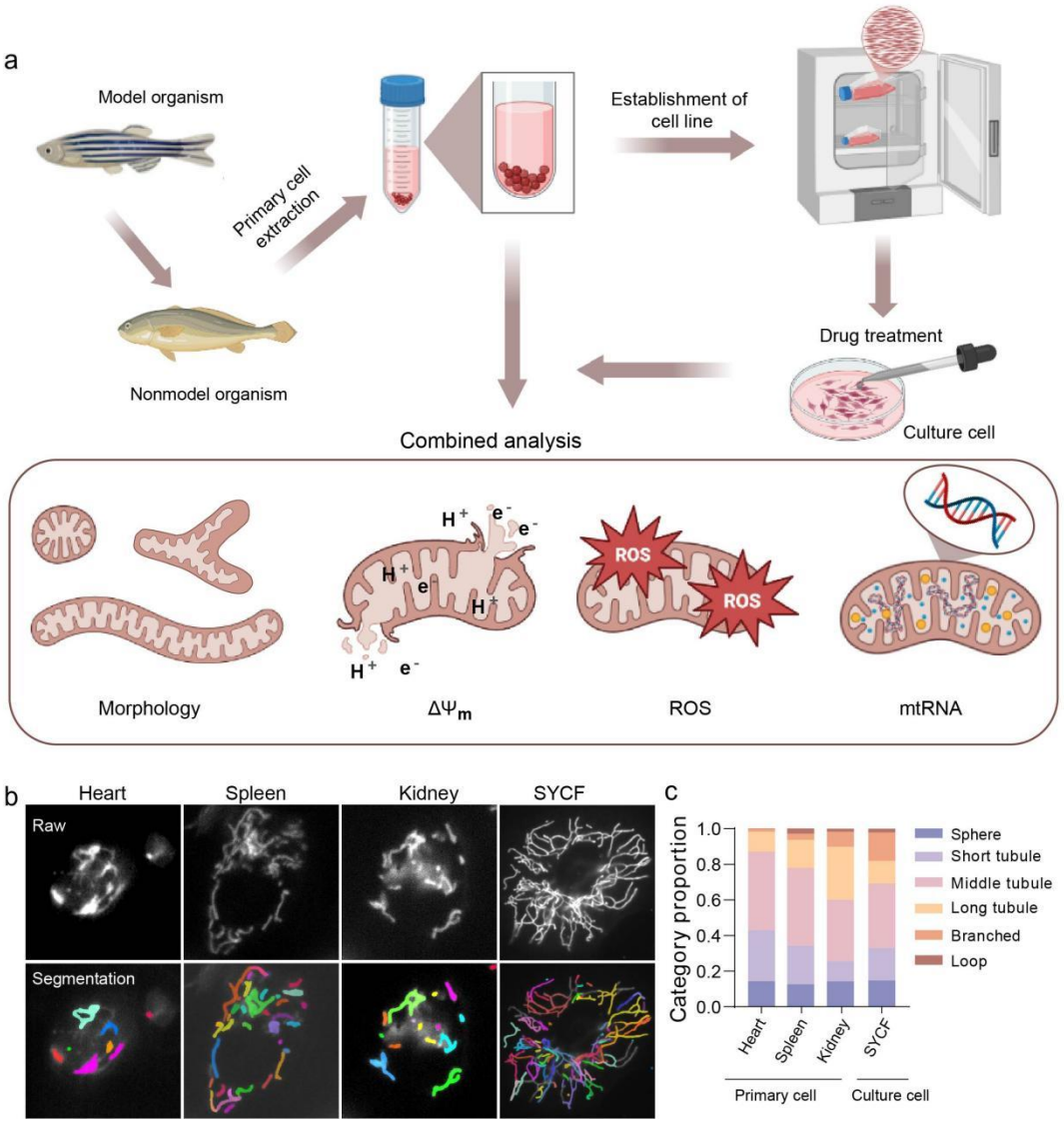


Fig. S15 – PanoMito enables adaptation of mitochondrial network analysis from model to non-model organism. (a) Workflow illustrating the adaptation of PanoMito-based mitochondrial network analysis from *Danio rerio* to non-model marine fish species. **(b)** Representative fluorescence and instance segmentation images of mitochondrial networks in primary and cultured cells derived from different tissues of *L. polyactis*. **(c)** Quantitative analysis of the six morphological categories in the primary and cultured cells from different tissues of *L. polyactis* shown in (b). Scale bar: 5 μm.

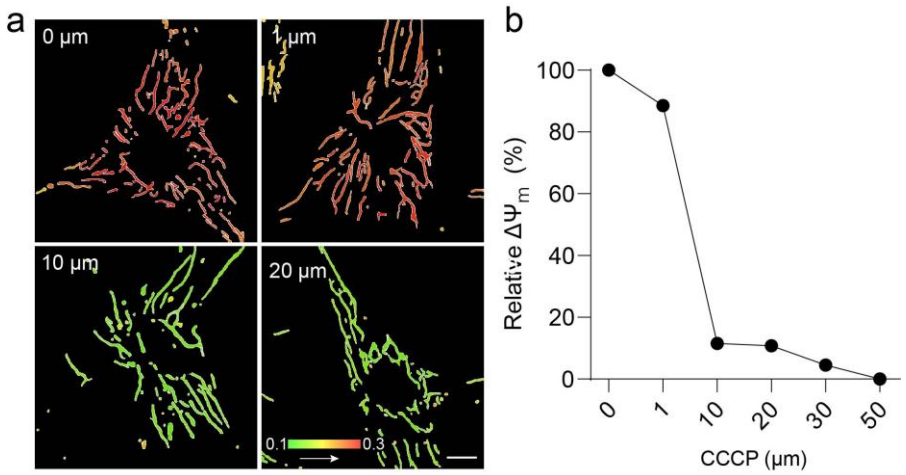


Fig. S16 – CCCP titration for TMRM response validation in SYCF cells. (a) Representative digital structured images of $\Delta\Psi_m$ in SYCF cells treated with different CCCP concentrations (0, 1, 10, 20, 30, 50 µM). Scale bar=5 µm. **(b)** Quantification of relative $\Delta\Psi_m$ (%) in SYCF cells treated with different CCCP concentrations.

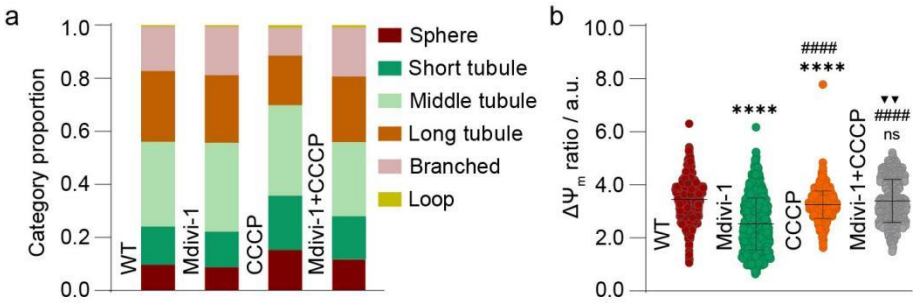


Fig. S17 – Quantitative analysis of mitochondrial morphology and $\Delta\Psi_m$ under pharmacological perturbations. (a) Proportion of six morphological categories under each pharmacological condition. **(b)** Comparison of $\Delta\Psi_m$ at the overall mitochondrion level across different treatments. Quantitative data are presented as the mean ± s.d. Significance ($p < 0.05$) for pairwise comparisons is denoted as follows: * vs the first group; # vs the second group; ▲ vs the third group.

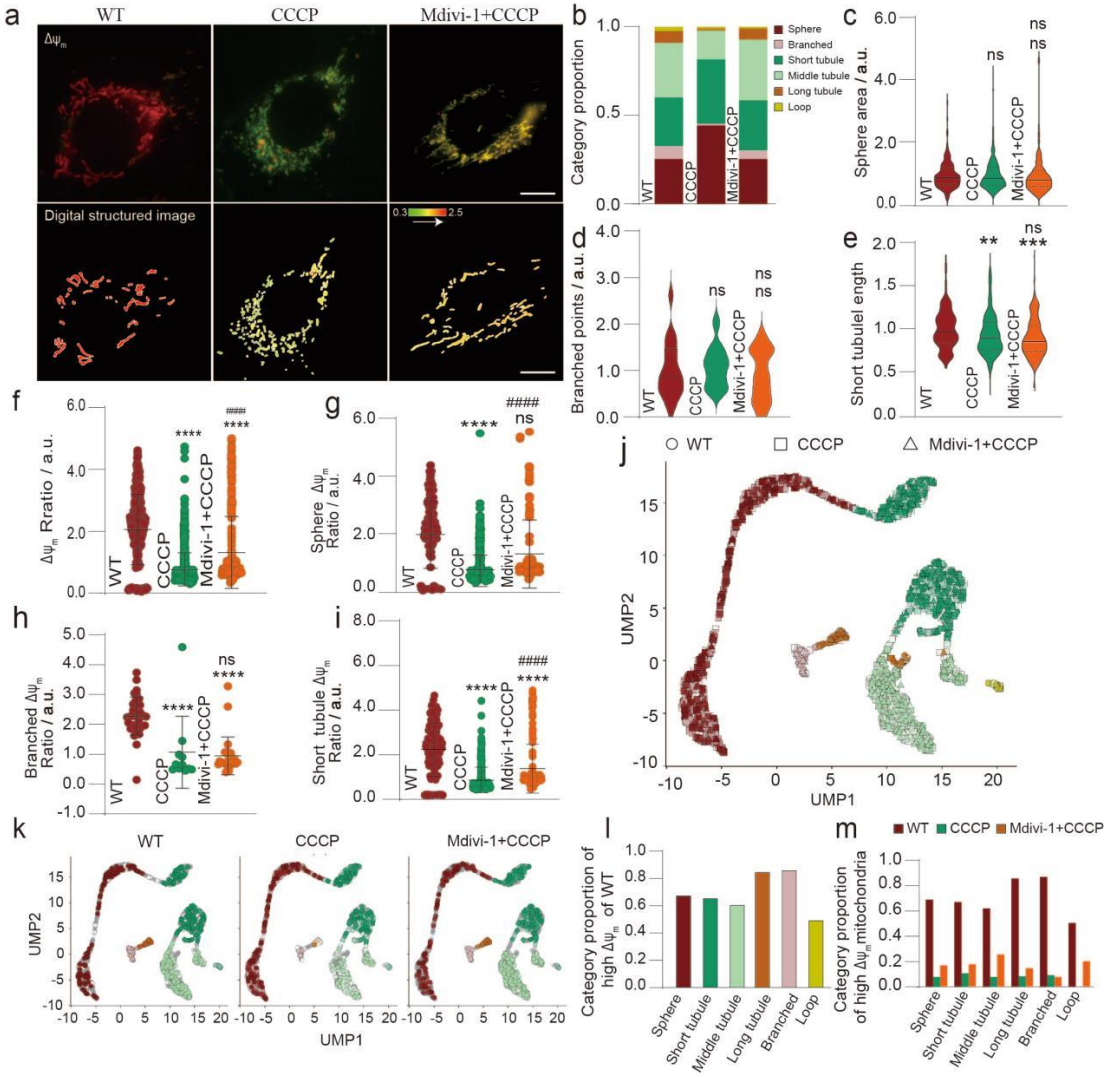


Fig. S18 – Intracellular mitochondrial heterogeneity of U2OS cells under pharmacological perturbations. (a) Representative $\Delta\Psi_m$ images and digital structural images of U2OS cells that were treated with CCCP or Mdivi-1 + CCCP. Scale bar=10 μm. **(b–e)** Quantitative analysis of mitochondrial morphology across treatment groups: (b) Proportion of six morphological categories. (c) Relative sphere area, (d) Relative branch points and (e) Relative short tubule length. **(f–i)** Quantitative comparison of $\Delta\Psi_m$ across treatments. $\Delta\Psi_m$ was compared (f) at the overall mitochondrial level, and within the subtypes of (g) sphere, (h) branched and (i) middle tubule. **(j)** UMAP projection of mitochondrial morphological clusters across treatment groups. Points denote single mitochondria, colored by their ratio value (darker = higher ratio). **(k)** UMAP projections for each treatment group, depicting the distribution of high $\Delta\Psi_m$ mitochondria. Mitochondria with a $\Delta\Psi_m \geq$ Q3 (the third quartile) are colored by morphological cluster; all others are shown in gray. **(l, m)** Proportion of high $\Delta\Psi_m$ mitochondria within each morphological cluster in the (l) WT group and (m) across treatment groups. Quantitative data are presented as the mean ± s.d. Significance ($p < 0.05$) for pairwise comparisons is denoted as follows: *, vs the first group; #, vs the second group; ▲, vs the third group.

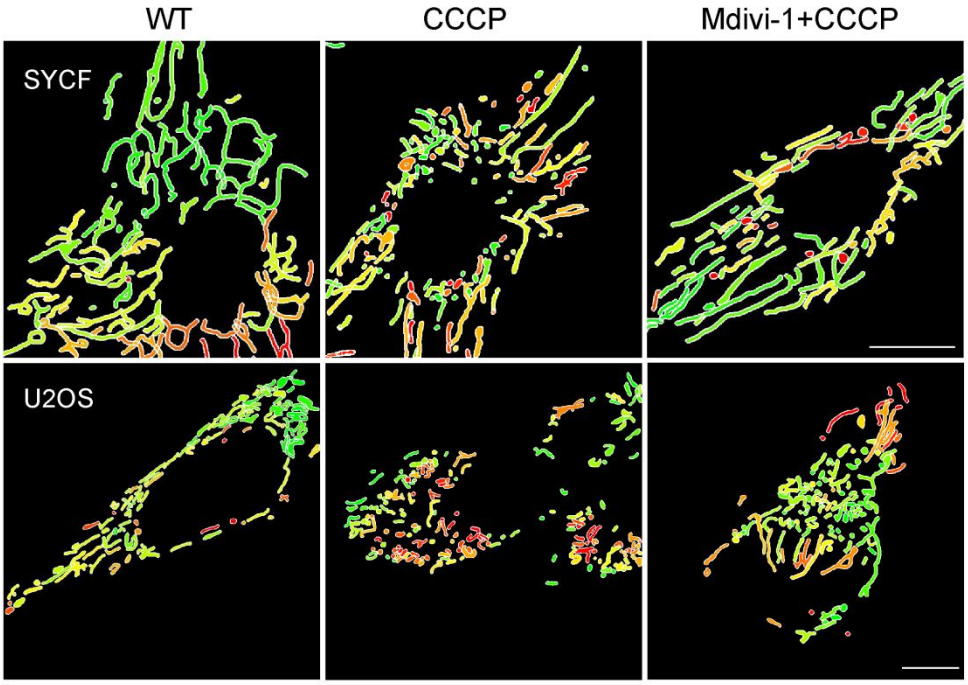


Fig. S19 – Representative structured images of drug-induced mtROS from the PanoMitoAtlas. Mitochondria are visualized with white outlines. Pseudocolor indicates instance wise average mtROS ratio in SYCF and U2OS cells under drug treatment, calculated as the MitoSOX/MitoTracker fluorescence ratio. Scale bar = 10 μm.

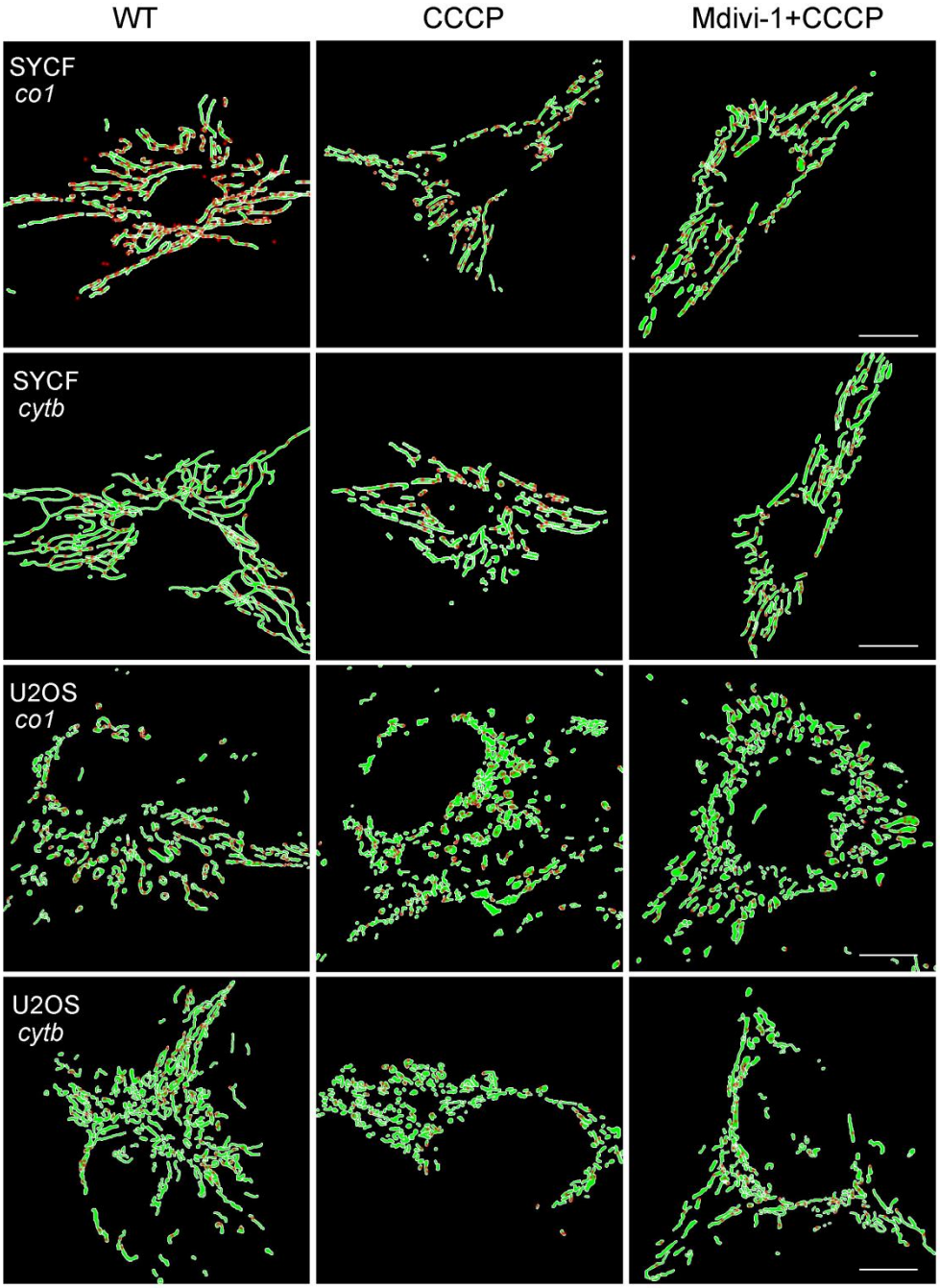


Fig. S20 – Representative structured images of mtRNA under pharmacological perturbation from the PanoMitoAtlas. Mitochondria are visualized with white outlines and green fill; mtRNA transcripts appear as red puncta for *cox1* or *cytb*. Scale bar = 10 μm.

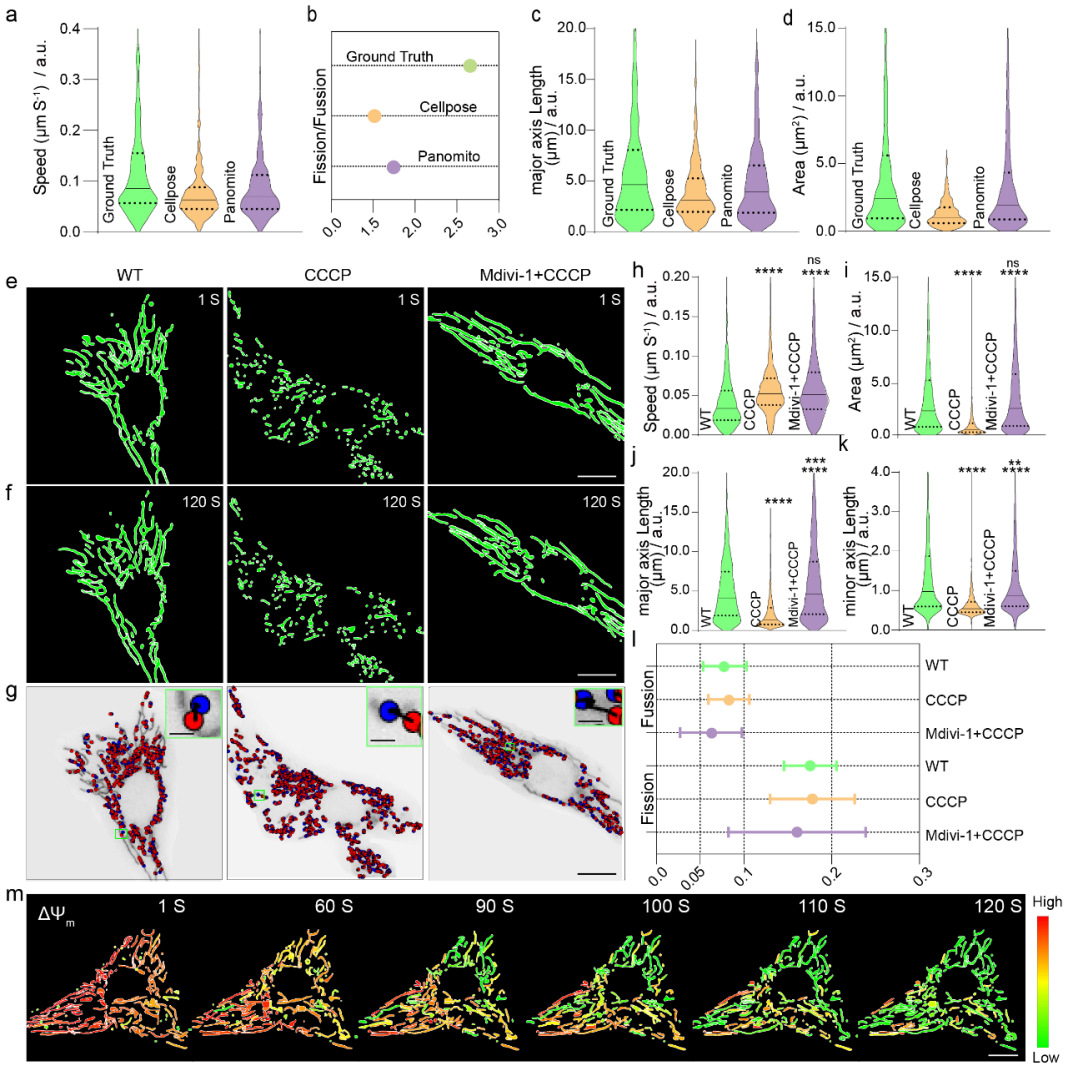


Fig. S21 – Mitochondrial dynamics revealed by temporal analysis. (a–d) Accuracy assessment of temporal analysis. Comparison of PanoMito against CellPose and manual annotations on a time-lapse imaging sequence of 30 images at 1 second interval in WT SYCF cells. Relative to ground-truth measurements, PanoMito yielded more accurate estimates than CellPose for (a) speed, (b) fission/fusion event rate, (c) major axis length, and (d) area. **(e–l)** Temporal dynamics under pharmacological perturbation. Analyses of SYCF cells treated with CCCP or Mdivi-1 + CCCP over 120 s at 1 second interval: (e) snapshot at 1 s timepoint, (f) snapshot at 120 s timepoint, and (g) merged tracking view over 120 s (the inset in the top right shows a magnified view of the region outlined in green, Scale bar = 1 μm). Quantitative analysis of (h) speed, (i) area, (j) major axis length, (k) minor axis length, and (l) fission-to-fusion rate. **(m)** Time-lapse structured images of SYCF cells during acute phototoxic stress over 120 s, where pseudocolor corresponds to $\Delta\Psi_m$. Scale bar = 10 μm. Significance ($p < 0.05$) for pairwise comparisons is denoted as follows: * vs the first group; # vs the second group; ▲ vs the third group.

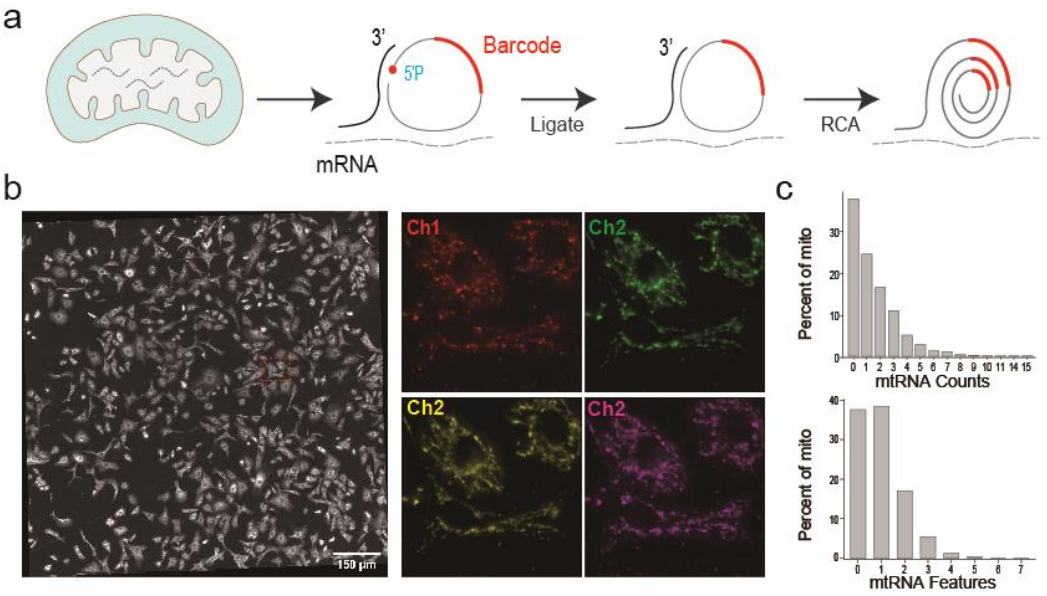


Fig. S22 – *In situ* sequencing of mitochondrial transcripts. (a) Schematic of probe-based *in situ* sequencing for mitochondrial mRNAs, involving hybridization, ligation, and cDNA amplicon generation. **(b)** Representative raw image showing four-color fluorescence channels (A, T, C, G) used for base decoding. **(c)** Distribution of mitochondrial transcriptome complexity. Upper: proportion of mitochondria binned by total mtRNA count; lower: proportion of mitochondria binned by number of distinct mtRNA species detected.

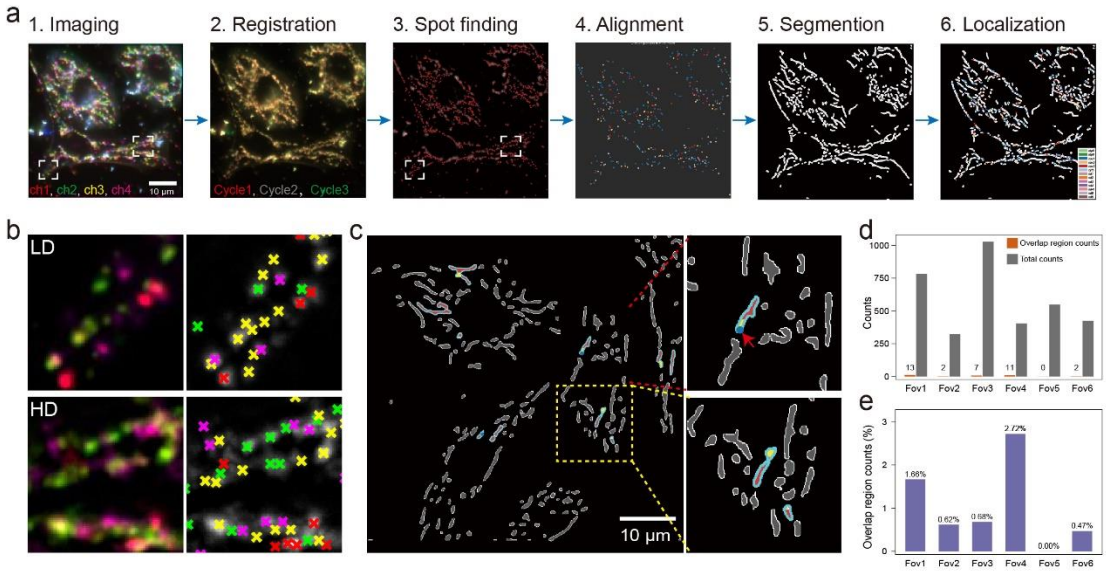
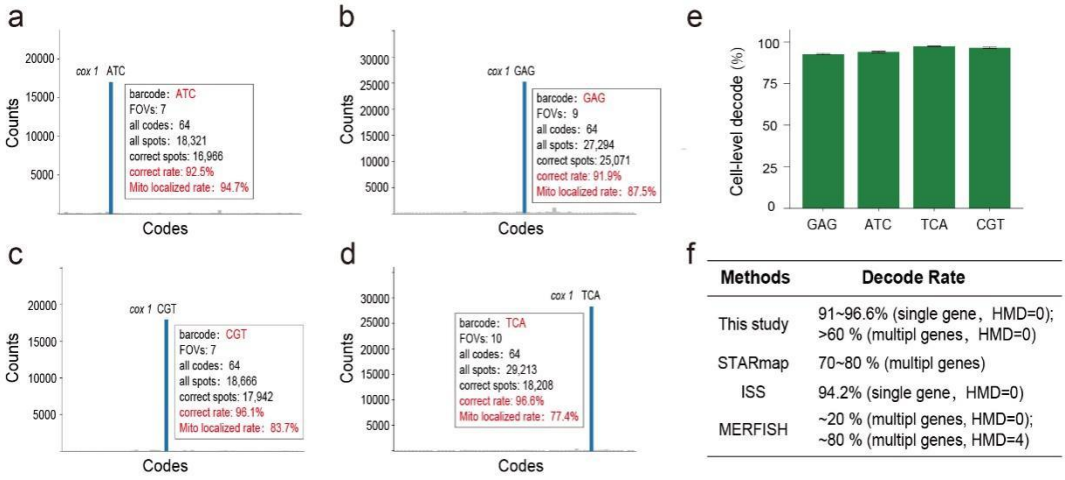


Fig. S23 – Assessment of transcript assignment ambiguity in densely packed mitochondrial regions. (a) Schematic overview of the transcript assignment workflow. Images from multiple sequencing cycles were first registered, followed by spot detection and barcode decoding through alignment to the predefined codebook. Mitochondria were subsequently segmented, and successfully decoded transcript spots were assigned to individual mitochondrial objects according to their spatial coordinates. **(b)** Fluorescent spot localization in high-density and low-density regions. (Top) Representative fluorescence image of a low-density region (left) and its corresponding schematic diagram illustrating the spatial distribution and centroid localization of individual fluorescent spots (right). (Bottom) Representative fluorescence image of a high-density region (left) and its corresponding schematic diagram illustrating the spatial distribution and centroid localization of individual fluorescent spots (right). **(c)** Representative fields of view showing segmented mitochondria and assigned transcript spots. The enlarged region (right) highlights a densely packed mitochondrial area containing spatial overlap between two neighboring segmented mitochondria (yellow region). Red arrows indicate transcript spots located within the overlapping region, which represent potentially ambiguous transcript assignments. **(d)** Quantification of total decoded transcript spots and transcript spots located within overlapping mitochondrial regions across six representative fields of view. **(e)** Fraction of transcript spots located within overlapping mitochondrial regions. Potentially ambiguous transcript assignments accounted for only 0–2.27% of all decoded transcript spots across analyzed fields of view.



Methods	Decode Rate
This study	91~96.6% (single gene, HMD=0); >60 % (multipl genes, HMD=0)
STARmap	70~80 % (multipl genes)
ISS	94.2% (single gene, HMD=0)
MERFISH	~20 % (multipl genes, HMD=0); ~80 % (multipl genes, HMD=4)

Fig. S24 – Validation of barcode decoding accuracy using four independent barcode designs targeting the same transcript. (a-d) Decoding performance of four independent 3-nt barcode designs targeting the same *cox1* transcript (ATC, GAG, CGT, and TCA, respectively). For each experiment, only the expected barcode was introduced, and all decoded barcode combinations were quantified. The dominant peak corresponds to the expected barcode, whereas alternative barcode assignments occurred at low frequencies. Correct decoding rates ranged from 91.9% to 96.6%, demonstrating high barcode assignment fidelity across barcode designs. Mitochondrial localization rates of correctly decoded puncta ranged from 77.4% to 94.7%. **(e)** Cell-level decoding performance for four independent 3-nt barcode designs targeting the same *cox1* transcript. Correct assignment rates were highly consistent across barcode designs (92.7–97.4%), indicating minimal nucleotide-specific or cycle-dependent decoding bias. **(f)** Summary of single-cell RNA sequencing and RCA-based multiplexed RNA detection methods, numbers were extracted from references[54-56].

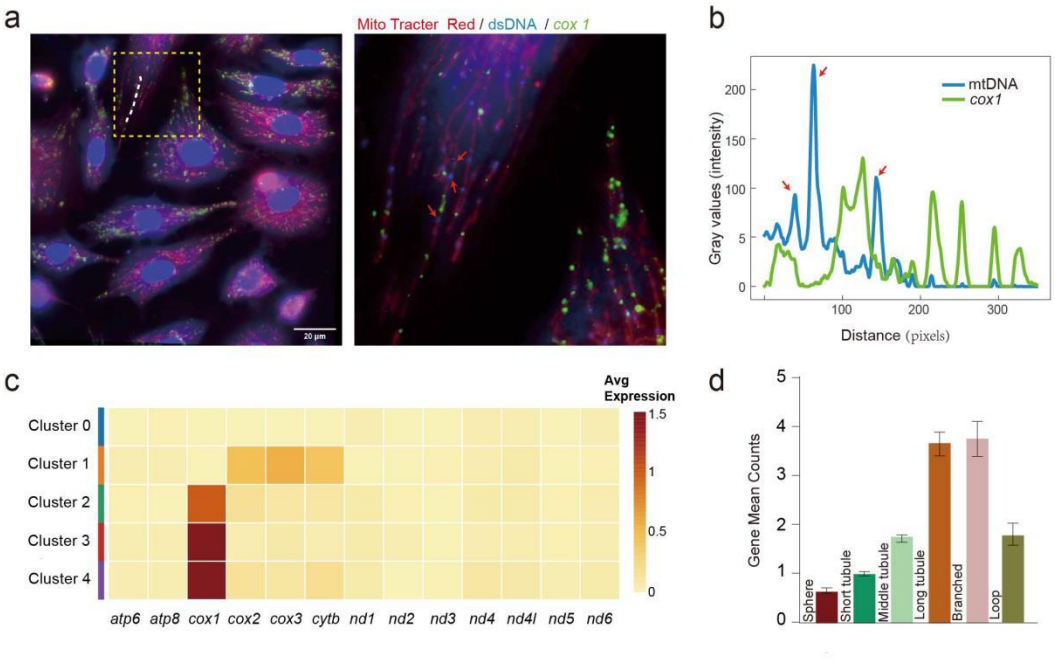


Fig. S25 – Simultaneous imaging of mitochondrial transcripts and mtDNA nucleoids, and morphology-associated mitochondrial transcriptomic states. (a) Representative fluorescence images showing mitochondria (red), mtDNA nucleoids (blue), and *cox1* mRNA transcripts (green). The merged image reveals the spatial distribution of *cox1* transcripts relative to mtDNA nucleoids within the mitochondrial network. Right-enlarged view of the boxed region. Arrows indicate representative mtDNA nucleoids. **(b)** Line-scan intensity analysis along the dashed line shown in **a**. Fluorescence intensity profiles of mtDNA (blue) and *cox1* mRNA (green) are plotted as a function of distance along the mitochondrial segment. Peaks indicate the locations of mtDNA nucleoids or *cox1* transcripts. While some *cox1* signals are observed adjacent to mtDNA nucleoids, many transcript signals are spatially separated from mtDNA peaks, indicating that mitochondrial transcripts are not exclusively confined to nucleoid-associated transcription sites. **(c)** Heatmap of differentially expressed mitochondrial genes across consolidated molecular states. Color intensity represents the mean expression level of each gene. **(d)** Mean mitochondrial gene expression across morphological categories.

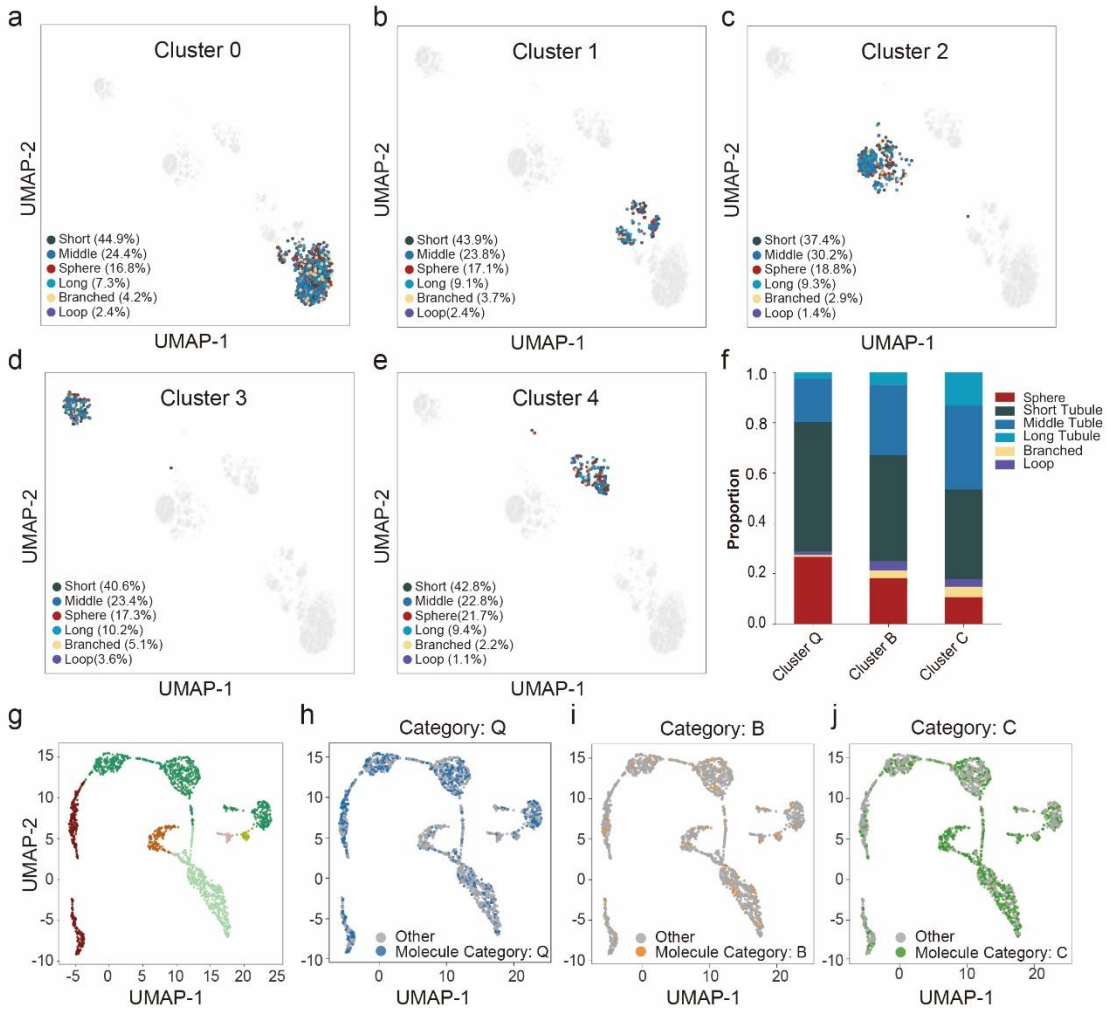


Fig. S26 – Association between molecular subtypes and morphology at single-mitochondrion resolution. (a-e) UMAP projection of single-mitochondria transcriptomes, stratified and colored by morphological category to show their distribution across five unsupervised transcriptomic clusters (remaining mitochondria in gray). **(f)** Compositional distribution of consolidated molecular subtypes across morphological categories. Original five unsupervised clusters were merged into three phenotypes (Cluster Q, Cluster B, and Cluster C) for biological interpretation. **(g-j)** UMAP projections of mitochondria from each morphological category, colored by assigned gene expression cluster to show their distribution across the three molecular phenotypes (remaining mitochondria in gray).

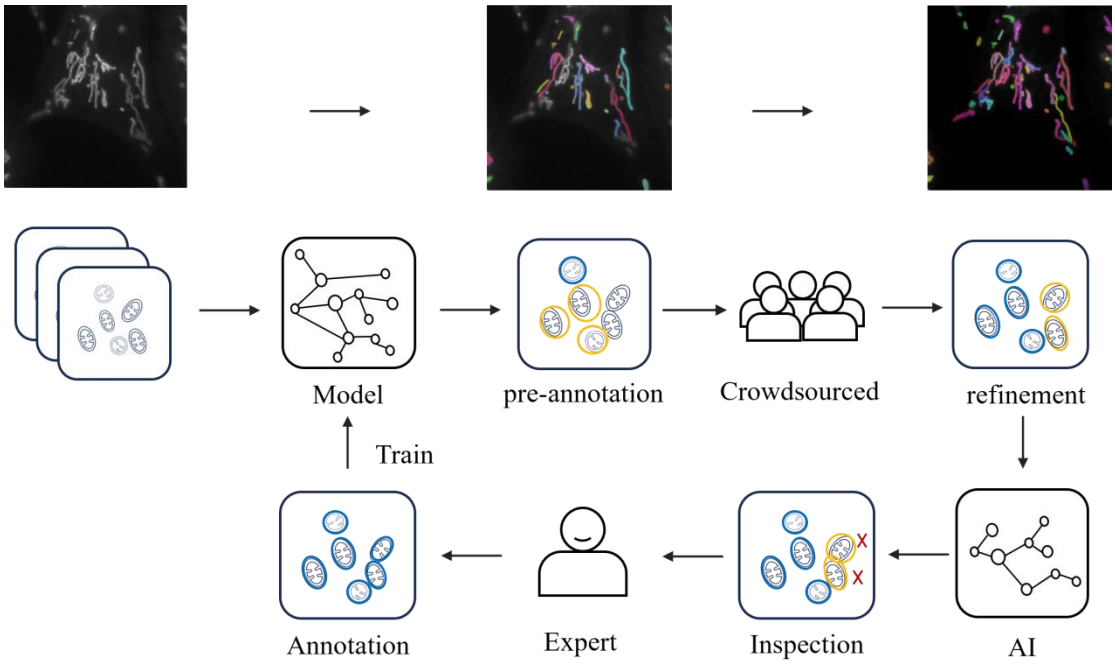


Fig. S27 – Iterative annotation workflow for constructing PanoMitoAtlas through iterative human-AI collaboration. Schematic of the human-AI collaborative annotation pipeline. An initial segmentation model was trained on a small set of manually annotated images to generate preliminary annotations. These pre-annotations were then manually corrected in CVAT by trained annotators following standardized criteria established to ensure consistency. An AI-based verification model subsequently performed instance-level quality control. Annotated mitochondrial objects were cropped and zoomed to 128×128 binary masks to normalize for size variation, then classified by a pretrained lightweight model. Certain categories directly indicated obvious errors—such as merged adjacent mitochondria, fragmented masks, or spurious small masks outside mitochondrial structures—which were flagged by AI-generated prompts for targeted expert correction to produce final ground-truth annotations. These high-quality annotations were incorporated into the training set to improve the segmentation model, which generated enhanced pre-annotations in the next cycle, progressively reducing manual effort. This iterative process continued until dataset completion.

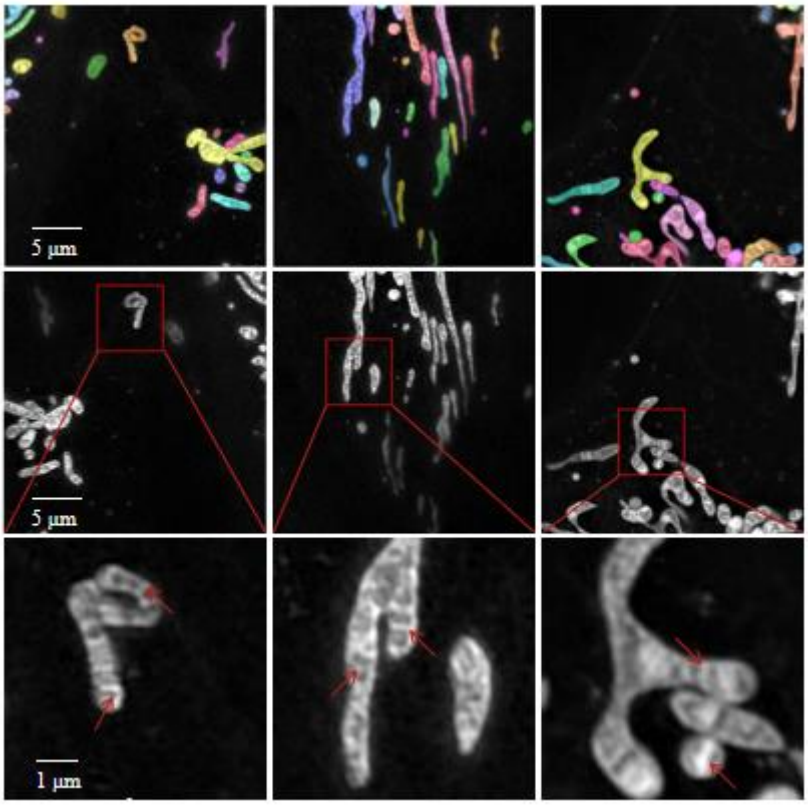


Fig. S28 – SIM imaging of mitochondria in U2OS cells. The top panel shows the PanoMitoNet segmentation result. The middle panel displays the original SIM image. The bottom panel presents a magnified view of the boxed region in the middle panel, with arrows indicating mitochondrial cristae.

Supplementary Tables

Table S1 – Composition of the PanoMitoAtlas dataset.

Type	Class	Species	Tissue	Function	Imaging	Treatment	Quantity
Primary Cell	Aves	*Columba livia domestica*	Spleen	$\Delta\Psi_m$	WF	WT	1679
	Reptilia	*Pelodiscus sinensis*	Kidney	$\Delta\Psi_m$	WF	WT	2523
	Amphibian	*Lithobates catesbeianus*	Liver	$\Delta\Psi_m$	WF	WT	1958
	Pisces	*Danio rerio*	Spleen	Intensity	WF	WT	1238
			Muscle	Intensity	WF	WT	518
			Ovary	Intensity	WF	WT	4358
			Intestines	Intensity	WF	WT	3565
			Kidney	Intensity	WF	WT	4955
		Lateolabrax japonicus	Blood	Intensity	WF	WT	1534
			Heart	Intensity	WF	WT	1677
			Ovary	Intensity	WF	WT	3012
		Tachysurus fulvidraco	Heart	$\Delta\Psi_m$	WF	WT	735
			Liver	$\Delta\Psi_m$	WF	WT	2246
			Kidney	$\Delta\Psi_m$	WF	WT	1277
			Spleen	$\Delta\Psi_m$	WF	WT	3744
		Larimichthys crocea	Ovary	Intensity	WF	WT	685
		Larimichthys polyactis	Kidney	$\Delta\Psi_m$	WF	WT	3595
	Crustacea	*Portunus trituberculatus*	Heart	$\Delta\Psi_m$	WF	WT	2673
			Ovary	$\Delta\Psi_m$	WF	WT	2649
		Litopenaeus vannamei	Gill	$\Delta\Psi_m$	WF	WT	3023
	Bivalvia	*Sinonovacula constricta*	Ovary	$\Delta\Psi_m$	WF	WT	5562
			Gill	$\Delta\Psi_m$	WF	WT	5252
	Insecta	*Drosophila melanogaster*	Worm	Intensity	WF	WT	2458
	Chlorophyta	*Chlamydomonas reinhardtii*	Worm	Intensity	WF	WT	862
Cultured Cell	Mammals	*Homo sapiens*_U2OS	Tibia	Intensity	Confocal	WT	12648
					SIM	WT	2015
					WF	WT	5437
						CCCP	6755
						Mdivi-1+CCCP	3289

		*Homo sapiens*_Hela	Uterus	Intensity	Confocal	WT	2024
		*Chlorocebus sabaeus*_COS7	Kidney	Intensity	Confocal	WT	2349
	Pisces	*Larimichthys polyactis*_SYCF	Fry	Intensity	WF	WT	5496
					WF	CCCP	3632
					WF	Mdivi-1+CCCP	2275
					WF	Hypoxia	2715
				$\Delta\Psi_m$	WF	WT	1371
					WF	CCCP	3140
					WF	Mdivi-1+CCCP	2156
				mtROS	WF	WT	483
					WF	CCCP	395
					WF	Mdivi-1+CCCP	174
				mtRNA	WF	WT	1085
					WF	CCCP	1077
					WF	Mdivi-1+CCCP	1294
				Time	WF	WT	858
				Protein	WF	WT	1551
Tissue	Mammals	*Mus musculus*	Articular cartilage	Intensity	STED	WT	8801
Total		132,798					

Note: This table includes detailed composition of the dataset, which includes the number of mitochondrial instances from different species, tissue types, sample types, diverse physiological/stress treatment and imaging modalities. With a total of 132,798 high-quality mitochondrial annotations, it establishes a comprehensive benchmark for evaluating segmentation algorithms across varied biological contexts and a systematic resource for interrogating mitochondrial structure–function relationships.

Table S2 – Implementation details of baseline methods and our proposed PanoMitoNet.

Model	Pretrained Model	Optimizer	Learning Rate	Epoch/ Iterations	Weight Decay	Batch Size
MitoMeter[17]	/	/	/	/	/	/
MoDL[15-16]	/	Adam	10^{-5}	100/-	0	2
MaskDINO[38]	TissueNet Pretrain	AdamW	10^{-5} 5000:10^{-6} 17000:10^{-7}	-/20000	0.05	16
CellPose-SAM[22]	CPSAM	AdamW	10^{-5}	500/-	0.1	1
OmniPose[23]	/	SGD	10^{-1}	4000/-	0.00001	16
PanoMitoSeg	TissueNet Pretrain	AdamW	0:10^{-5} 5000:10^{-6} 17000:10^{-7}	-/20000	0.05	16

Note: This table summarizes the key hyperparameters and training configurations for all methods compared in the study. MitoMeter is a traditional algorithm without training settings. MoDL was retrained on PanoMitoAtlas, and the listed setting was selected based on parameter tuning. CellPose-SAM was fine-tuned from the CPSAM pretrained model, and OmniPose was trained using the authors' recommended settings. MaskDINO and our PanoMitoNet employs a multi-phase learning rate schedule (1e-5, 1e-6, 1e-7) with linear warm-up and cosine decay, pretrained on the TissueNet dataset. All experiments were conducted under consistent hardware and software environments to ensure fair comparison.

Table S3 – Parameter-tuning results for benchmarked methods.

Method	Train epoch	Prob threshold	AP50	F1score	Accuracy
MoDL	50	0.8	46.5	42.5	27.0
		0.7	51.5	50.1	33.4
		0.6	51.5	52.4	35.5
		0.5	49.5	52.4	35.5
		0.4	45.5	50.2	33.5
		0.3	39.6	45.7	29.7
	100	0.9	51.5	46.3	30.1
		0.8	54.5	52.8	35.9
		0.7	52.5	52.6	35.7
		0.6	48.5	50.1	33.7
		0.5	44.6	46.4	30.2

Method	Cellprob	Flow	Diameters	AP50	F1score	Accuracy
CellPose	-1.5	0.3	30	38.6	52.0	35.2
			25	40.6	54.1	37.1
			20	41.6	52.7	35.8
		0.4	30	52.5	60.8	43.7
			25	55.4	61.8	44.7
			20	55.4	57.8	40.7
		0.5	30	61.4	62.5	45.5
			25	62.4	62.6	45.6
			20	63.4	56.9	39.8
	-1.0	0.3	30	45.5	57.6	40.4
			25	47.5	59.3	42.2
			20	48.5	57.1	40.0
		0.4	30	57.4	63.4	46.5
			25	59.4	64.1	47.1
			20	59.4	59.8	42.7
		0.5	30	64.4	63.8	46.8
			25	65.3	63.5	46.5
			20	65.3	56.9	39.8
	-0.5	0.3	30	49.5	60.2	43.1
			25	51.5	61.5	44.4
			20	52.5	59.0	41.8
		0.4	30	59.4	63.7	46.8
			25	61.4	63.9	46.9
			20	60.4	59.0	41.9
		0.5	30	63.4	63.1	46.1
			25	65.3	61.7	44.6
			20	64.4	54.4	37.4
	0.0	0.3	30	50.5	59.8	42.6
			25	52.5	60.8	43.7

		20	52.5	58.3	41.1
		30	57.4	61.3	44.2
	0.4	25	59.4	61.3	44.2
		20	58.4	57.1	40.0
		30	60.4	60.4	43.2
	0.5	25	61.4	58.0	40.9
		20	60.4	52.8	35.8
		30	47.5	56.8	39.6
	0.3	25	50.5	58.0	40.8
		20	49.5	55.7	38.6
		30	53.5	57.2	40.1
0.5	0.4	25	55.4	57.6	40.4
		20	53.5	54.1	37.1
		30	55.4	56.0	38.9
	0.5	25	56.4	54.9	37.8
		20	55.4	51.6	34.7
		30	42.6	51.8	35.0
	0.3	25	45.5	53.1	36.2
		20	44.6	50.8	34.0
		30	46.5	51.7	34.9
1.0	0.4	25	48.5	52.6	35.7
		20	46.5	49.6	33.0
		30	48.5	50.2	33.6
	0.5	25	50.5	51.0	34.2
		20	48.5	48.0	31.6

Note: For MoDL, different training epochs and probability thresholds were evaluated. For CellPose-SAM, different cell probability thresholds, flow thresholds, and object diameters were tested. Performance was assessed using AP50, F1-score, and instance detection accuracy. The parameter settings highlighted in red indicate the final configurations used for the benchmark comparison in the Results

Table S4 – Comparison of mitochondria specific and generalist segmentation methods by application scope, instance-segmentation capability, approach, downstream analysis, and key limitations.

Method	Generalist / Mito-specific	Instance segmentation	Segmentation approach	Downstream analysis	Key limitation
PanoMitoSeg	Mito-specific	Yes	End-to-end instance segmentation	Morphology clustering + Morphology-function-molecular profiling	Currently focuses mainly on 2D static segmentation
MitoHacker[24]	Mito-specific	No	Semantic segmentation	Cell-level and mitochondria-level morphometric analysis	Does not predict individual mitochondrial instance masks
Momito[25]	Mito-specific	No	Post-segmentation skeleton/topology analysis	None	Requires pre-segmented binary mitochondrial images as input
MitoSegNet[26]	Mito-specific	No	Semantic segmentation	Mitochondrial morphology quantification from pixel-level masks	Produces semantic/binary mitochondrial masks rather than individual instance masks
MitoMeter[17]	Mito-specific	Yes	Thresholding-based segmentation	Segmentation, tracking, and quantification of mitochondrial morphology	Limited performance across heterogeneous modalities and species.
MoDL[15,16]	Mito-specific	Yes	Two-step instance segmentation	Links segmented mitochondrial morphology to image-level functional states	Relies on semantic mask prediction followed by connected-component post-processing
CellPose[19-22]	General	Yes	Two-step instance segmentation	None	Generalist model not optimized for filamentous, branched, and densely adherent mitochondrial networks
OmniPose[23]	Bacteria-specific	Yes	Two-step instance segmentation	Bacterial morphology analysis	Optimized for bacterial cell segmentation rather than mitochondria segmentation
MaskDINO[38]	General	Yes	End-to-end instance segmentation	None	General-purpose instance segmentation framework without mitochondria-specific continuity or topology priors

Table S5 - The primer sequences used in this study.

Species	Name	Sequence
Danio rerio	*nd4*_01	ACATTATCCGAGGTTCATTTTAGTCACAACTCCTTGGCTCACAGA ACGACATGACAATGCTCAAGATA
	*nd4*_02	GTTAGAAGCATTTCACCCTGTAATGTTATCTT
*Larimichthys polyactis*_SYCF	*co1*_01	ACATTATGCGAGATTTCCAGCAAGTGCAACTCCTTGGCTCACAG AACGACATGAGTCAGCTGAAGATA
	*co1*_02	GTCGACTGAAGCTCCTGCGTTAATGTTATCTT
	*cytb*_01	ACATTACGAATAGGAAGTATCACTCTGGCTCAACTCCTTGGCTC ACAGAACGACATTGTCTGCGAAAAGATA
	*cytb*_02	TCGATCGGAGGATGGCGTAGTAATGTTATCTT
	*co2*_01	ACATTACAGGGTGGGAACTGCTCATCAACTCCTTGGCTCACAGAA CGACATTTGCCATCTCAAGATA
	*co2*_02	ACTGCGTCTATTTTGATGCTAATGTTATCTT
	*co3*_01	ACATTAAGGACGGCAGTATTGAGCACAACTCCTTGGCTCACAGA ACGACATCCACTAGTCCAAGATA
	*co3*_02	AACTGTTACACCAGAGGCTATAATGTTATCTT
	*atp6*_01	ACATTAAGAACTATCGTAAGAGCCGCCAACTCCTTGGCTCACAGA ACGACATGTCCAGAGCTAAGATA
	*atp6*_02	TCAAGTAGGGTGAGTAGGAATAGTAATGTTATCTT
	*atp8*_01	ACATTATTGTGGGGAGGGTTCGTTCAACTCCTTGGCTCACAGAAC GACATCGCTATCGGCAAGATA
	*atp8*_02	TCTATGTCTAAGGCTTCTGTGCTAATGTTATCTT
	*nd1*_01	ACATTAGAGAAGGGCTGCTTTGGTCAACTCCTTGGCTCACAGAA CGACATGGTTCCACACAAGATA
	*nd1*_02	GGACTCATAGGAAGAGGACAGTAATGTTATCTT
	*nd2*_01	ACATTAAGTGTAACGGGGAGTGGGTCAACTCCTTGGCTCACAGA ACGACATGCTGAGCTGTAAGATA
	*nd2*_02	GAGCGAGGGCGATGGTGAATAATGTTATCTT
	*nd3*_01	ACATTAGGGAGTCAGAATGATACAGCAACTCCTTGGCTCACAGA ACGACATCCGCTCAGTAAAGATA
	*nd3*_02	TTCTCGTGGTCAGGATTTATTAATGTTATCTT
	*nd4l*_01	ACATTATCCGTGTGTTCGGGTAGTGGCAACTCCTTGGCTCACAGA

		ACGACATAGTTGCCATAAAGATA
	*nd4l*_02	GCTTTGAAGGCGGTCGGTAATGTTATCTT
	*nd4*_01	ACATTAGGTCAAAGTCATTTAGCGGGCAACTCCTTGGCTCACAGA ACGACATATGTACGCAGAAGATA
	*nd4*_02	TAGGCTGTGCATAAGGGTGGTAATGTTATCTT
	*nd5*_01	ACATTACAGCCGATTAGGAGGAAAGACAACTCCTTGGCTCACAG AACGACATTGCGCCACTTAAGATA
	*nd5*_02	CATCTGCTCGCCCGTATCATAATGTTATCTT
	*nd6*_01	ACATTACAAGAGGTCTCATACCATCCCCAACTCCTTGGCTCACAG AACGACATAATCGGCTCGAAGATA
	*nd6*_02	CCCAACTCATCTACAGGCACTAATGTTATCTT