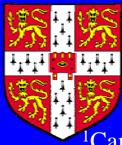


High-throughput analysis of protein localization dynamics by mass spectrometry

Denise J L Tan[1,2], **Svenja Hester**[1], **Alfonso Martinez-Arias**[2] and **Kathryn Lilley**[1]

[1]Cambridge Systems Biology Centre, Dept of Biochemistry, Tennis Court Road, Cambridge, CB2 1QR, UK. [2] Dept of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH.

Abstract

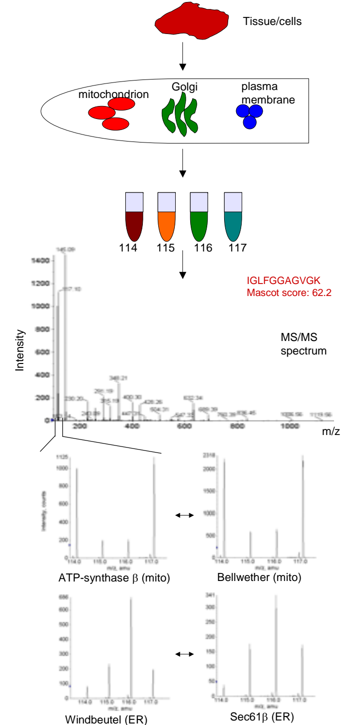
While the subcellular localization of a protein provides one with an initial clue to its cellular function, dynamic changes in protein localization are even more informative as they are often critical for regulation of protein activity. Previously, we had developed LOPIT, a high-throughput technique of localizing proteins to subcellular organelles. Here, we demonstrate how this technique was applied to assign proteins in *Drosophila melanogaster* embryos and propose how it may be extended to allow analysis of dynamic changes in protein localization during signaling - Dynamic LOPIT.

1) Localization of Proteins by Isotope Tagging (LOPIT)

• The living cell is organised spatially and functionally into specialized organelles containing proteins. Hence the subcellular localization of a protein provides an important clue to its function.

• Localization of Proteins by Isotope Tagging (LOPIT) is a high-throughput technique of protein localization by mass spectrometry (Dunkley TP *et al*. PNAS 103(17):6518-23).

• Organelles are first partially separated on density gradients. Protein distributions within the gradient are determined by labeling gradient fractions with isotope-coded iTRAQ reagents. Proteins from the same organelle show similar distributions in the gradient, so unknown proteins can be localized by comparing their distributions to that of known organelle markers.



1) Lysis

2) Density gradient centrifugation for partial organelle separation

3) Fractions collected and proteins digested to peptides. Peptides from each fraction labelled with a different isotopic iTRAQ tag for quantitation.

4) Quantitation and Identification. Peptides analyzed by 2D-LC-MS/MS. In MS/MS mode, fragmentation of peptide ions give rise to iTRAQ reporter peaks for quantitation.

5) The reporter ion ratios indicate a given protein's distribution in the gradient. Proteins from the same organelle share similar distributions. Hence unknowns can be assigned to organelles by comparing their distributions to those of known markers.

6) Data is analyzed by principal components analysis (PCA), which clusters proteins according to their co-variation.

Fig. 1. LOPIT workflow

2) LOPIT successfully applied to localize *Drosophila* embryos proteins

LOPIT was initially developed for protein localization in *Arabidopsis* tissue culture. In this study, we have applied the technique to studies in a whole organism, the *Drosophila* embryo.

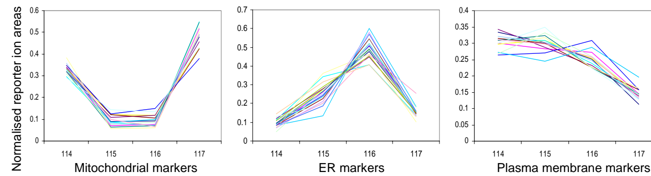


Fig. 2. Marker proteins from the same organelle show similar distributions, while proteins from different organelles exhibit distinct distributions.

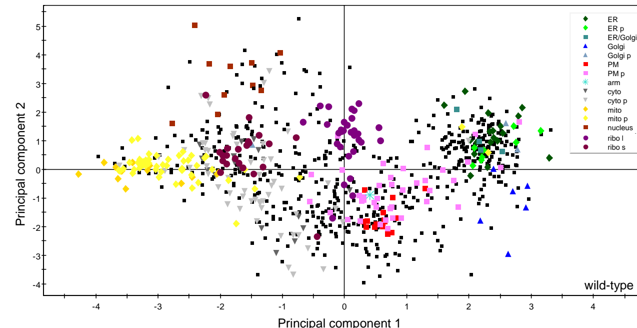


Fig. 3. Principal components analysis (PCA) scatter plot showing organelle clusters. PCA was used to obtain an overview of the dataset. Proteins with similar distributions cluster together. Unknown proteins that cluster together with known organelle markers can be assigned putative localizations. Novel complexes can be identified as a cloud of points eg. ribosome.

Protein ID	*LOPIT Localization*	*Validation*
CG32675	ER	Identified in an RNAi screen for genes involved in secretion. V5-tagged CG32675 (red) shows a characteristic perinuclear ER localization. (Figure adapted from Bard *et al*., 2006, Nature 439: 604 with permission from Macmillan Publishers Ltd.)
CG33303	ER	Identified in RNAi screen as above.
CG6370	ER	Identified in RNAi screen as above.
CG10960	Plasma membrane	Predicted sugar transporter, 12 predicted TM helices (TMHMM 2.0)
CG31729	Plasma membrane	Sequence similarity to mouse lipid transporter, 9 predicted TMH
Mipp1	Plasma membrane	Homologue of human multiple inositol polyphospholipid phosphatase1
CG7998	mitochondrion	Identified in MitoDrome, database of *Drosophila* mitochondrial proteins identified by sequence similarity to human (D'Elia *et al*., 2006, Mitochondrion 6:252).
CG4692	mitochondrion	Identified in MitoDrome as above.
CG6459	mitochondrion	Identified in MitoDrome as above.

Table 1. Examples of proteins classified into organelles by LOPIT. To classify proteins into organelles, partial least squares-discriminant analysis (PLS-DA) was used to build models of organelles using a training set of at least 15 marker proteins per organelle. Unknown proteins were then tested against this model to obtain confident classifications.

3) Extending LOPIT to analysis of protein dynamics

• The power of LOPIT lies in its ability to distinguish multiple organelles *simultaneously*. Hence it makes it possible to study global changes in protein localization. Changes in localization are often involved in regulation of protein activity, particularly in signal transduction pathways.

• In this study, the Wnt/Wingless signaling pathway in *Drosophila* was chosen as a model system. Several components of the pathway are known to move during signaling, hence studying global changes in localization will provide new insights into its mechanism.

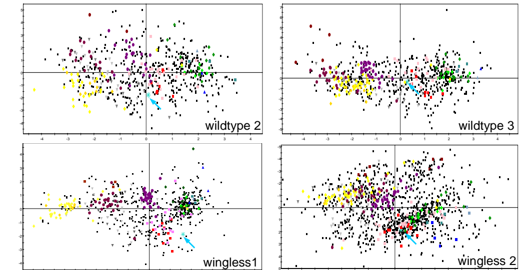


Fig. 4. Dynamic LOPIT applied to wild-type and wingless-activated* Drosophila *embryos. LOPIT analysis was carried out in wild-type and wingless-activated embryos to identify proteins that alter their steady-state location due to signaling. Preliminary data indicated that this approach could identify moving proteins, however further repeats revealed that the variation was problematic (see blue arrow). Classification of proteins within organelles was reproducible, however absolute positions between experiments was variable, making it difficult to compare data across experiments.

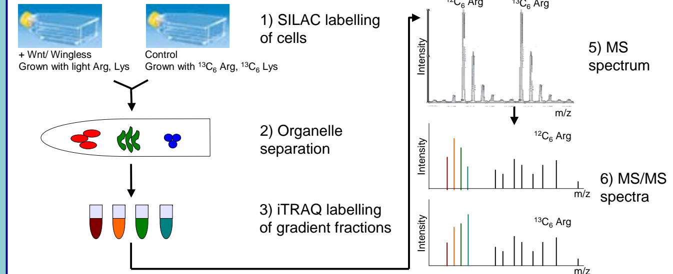


Fig. 5. Proposed Dynamic LOPIT workflow. Combining SILAC (to distinguish biological samples) with iTRAQ (to distinguish gradient fractions) may overcome the problem of variation between LOPIT experiments, hence allowing one to identify proteins that are truly moving.

4) Conclusions

• High-throughput protein localization by LOPIT was applied to a whole organism- the *Drosophila* embryo. 851 proteins were identified and many organelles could be distinguished.

• Novel proteins can be assigned to subcellular organelles, providing an initial clue to protein function, or validating annotations made by sequence homology.

• The simultaneous localization of proteins to multiple organelles provides the potential for studies of protein dynamics in response to cellular perturbations.

Acknowledgements

We would like to thank the BBSRC for funding this work, Tina Balayo for help with growing flies, Jules Griffin, Paul Bertone and Heidi Dvinge for help with statistics, and members of the CCP and AMA labs for general advice and discussions.