Proteomic dissection of the Arabidopsis Golgi and trans-Golgi network

Harriet T. Parsons1*, Georgia Drakakaki2 and Joshua L. Heazlewood3,4

1 Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark
2 Department of Plant Sciences, University of California at Davis, Davis, CA, USA
3 Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
4 Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

*Correspondence: Harriet T. Parsons, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensgade 40, 1871 Frederiksberg C, Copenhagen, Denmark. e-mail: htpa@life.ku.dk

BACKGROUND

At its simplest level, subcellular proteomics attempts to identify all proteins in a particular compartment. However, even with such a basic definition in mind, the Golgi proteome presents conceptual difficulties; functional proteins in the Golgi may also be functional elsewhere (Ondzighi et al., 2008), whilst endoplasmic reticulum (ER)-Golgi connections (Boevink et al., 1998) makes absolute divisions between the proteomes of these compartments somewhat futile. A number of proteins are known to form functional associations on the cytoplasmic face of cisternae but are part of the cytosol (Ito et al., 2011), so the very definition of the Golgi proteome is problematic. Furthermore, in such an architecturally heterogeneous organelle, simply identifying all the proteins present in the Golgi is not that helpful unless we can classify them according to sub-Golgi location, post-Golgi compartments, cargo, resident, or dual-localized proteins. The plant Golgi apparatus and trans-Golgi network (TGN) are major endomembrane trafficking hubs within the plant cell and are involved in a diverse and vital series of functions to maintain plant growth and development. Recently, a series of disparate technical approaches have been used to isolate and characterize components of these complex organelles by mass spectrometry in the model plant Arabidopsis thaliana. Collectively, these studies have increased the number of Golgi and vesicular localized proteins identified by mass spectrometry to nearly 500 proteins. We have sought to provide a brief overview of these technical approaches and bring the datasets together to examine how they can reveal insights into the secretory pathway.

Keywords: Golgi, trans-Golgi network, proteomics, LOPIT, free-flow electrophoresis, Arabidopsis, SYPER1

The plant Golgi apparatus and trans-Golgi network are major endomembrane trafficking hubs within the plant cell and are involved in a diverse and vital series of functions to maintain plant growth and development. Recently, a series of disparate technical approaches have been used to isolate and characterize components of these complex organelles by mass spectrometry in the model plant Arabidopsis thaliana. Collectively, these studies have increased the number of Golgi and vesicular localized proteins identified by mass spectrometry to nearly 500 proteins. We have sought to provide a brief overview of these technical approaches and bring the datasets together to examine how they can reveal insights into the secretory pathway.

Edited by: Rebecca L. Taylor, The University of Western Australia, Australia
Reviewed by: Holger Eubel, Leibniz Universität Hannover, Germany
Karina Gallardo, National Institute for Agrobiomass Research, France

The plant Golgi apparatus and trans-Golgi network (TGN) are major endomembrane trafficking hubs within the plant cell and are involved in a diverse and vital series of functions to maintain plant growth and development. Recently, a series of disparate technical approaches have been used to isolate and characterize components of these complex organelles by mass spectrometry in the model plant Arabidopsis thaliana. Collectively, these studies have increased the number of Golgi and vesicular localized proteins identified by mass spectrometry to nearly 500 proteins. We have sought to provide a brief overview of these technical approaches and bring the datasets together to examine how they can reveal insights into the secretory pathway.

Keywords: Golgi, trans-Golgi network, proteomics, LOPIT, free-flow electrophoresis, Arabidopsis, SYPER1

The plant Golgi apparatus and trans-Golgi network (TGN) are major endomembrane trafficking hubs within the plant cell and are involved in a diverse and vital series of functions to maintain plant growth and development. Recently, a series of disparate technical approaches have been used to isolate and characterize components of these complex organelles by mass spectrometry in the model plant Arabidopsis thaliana. Collectively, these studies have increased the number of Golgi and vesicular localized proteins identified by mass spectrometry to nearly 500 proteins. We have sought to provide a brief overview of these technical approaches and bring the datasets together to examine how they can reveal insights into the secretory pathway.

Keywords: Golgi, trans-Golgi network, proteomics, LOPIT, free-flow electrophoresis, Arabidopsis, SYPER1

The plant Golgi apparatus and trans-Golgi network (TGN) are major endomembrane trafficking hubs within the plant cell and are involved in a diverse and vital series of functions to maintain plant growth and development. Recently, a series of disparate technical approaches have been used to isolate and characterize components of these complex organelles by mass spectrometry in the model plant Arabidopsis thaliana. Collectively, these studies have increased the number of Golgi and vesicular localized proteins identified by mass spectrometry to nearly 500 proteins. We have sought to provide a brief overview of these technical approaches and bring the datasets together to examine how they can reveal insights into the secretory pathway.

Keywords: Golgi, trans-Golgi network, proteomics, LOPIT, free-flow electrophoresis, Arabidopsis, SYPER1
during partial least squares discriminant analysis (Figure 1). Using LOPIT, 89 proteins were initially localized to the Golgi (Dunkley et al., 2006) but the requirement that proteins carry all four tags limited the number of proteins for which a statistically credible localization could be assigned. Recent reanalysis and analysis of existing and new datasets, incorporating values for “missing” tags assigned using partial least squares regression models and training sets based on fully tagged proteins, enabled the collective localization of 204 proteins to the Golgi (Dunkley et al., 2006; Nikolovski et al., 2012).

Although a major motivation for the development of LOPIT was the difficulty in separating the Golgi, particularly from ER contaminants, a recent study has managed to isolate Golgi vesicles with an estimated 80% purity based on protein composition. This was achieved using a combination of sucrose density centrifugation and FFE (Parsons et al., 2012a). The power of FFE for organelle isolation was demonstrated in plants several years ago when applied to the separation of mitochondria and peroxisomes, two organelles which are typically hard to separate using density centrifugation alone (Eubel et al., 2008). As separation by FFE is dependent on surface charge, the Golgi, which carries a more negative surface charge than ER vesicles and most other contaminants, is amenable to separation using this technique, which resulted in 371 proteins being localized to the Golgi (Figure 1).

FIGURE 1 | Overview of the three different techniques employed in proteomic characterization of the Arabidopsis Golgi and TGN. (A) Clustered proteins in LOPIT studies were assigned to the Golgi according to co-clustering with known and predicted Golgi marker proteins. For details, see Dunkley et al., 2004, 2006; Nikolovski et al., 2012. (B) FFE purified fractions were estimated at ca. 80% purity according to the proportion of previously localized Golgi proteins and contaminants present in each fraction; based on experimental data in SUBA (Heazlewood et al., 2007; for details, see Parsons et al., 2012a; b). (C) Isolation of SYP61 vesicles by affinity purification. Successful removal of contaminants during immunoisolation was assayed by the presence of the ER/cis-Golgi marker, BiP, and the prevacuolar compartment marker SYP21 (for details, see Drakakaki et al., 2012; Parsons et al., 2013a).
A number of large gene families have been identified by both the AdaBoost, Golgi and PProwler, Golgi/TGN proteome is estimated to be 2239. Arabidopsis found by both FFE and LOPIT approaches (Nikolovski et al., 2012; Chou and Gasser, 1997). In plants, they are classically associated with the thylakoid lumen where they are thought to help protein folding and assembly of photosystem complexes although their exact role is not clear (Ingelsson et al., 2009). The cyclophilins found by both FFE and LOPIT approaches (Nikolovski et al., 2012; Parsons et al., 2012a) localize either exclusively to the Golgi or are dually localized to the Golgi and plasma membrane (Dunkley et al., 2006; Benschop et al., 2007; Marmagne et al., 2007; Parsons et al., 2012a), implying a secretory-specific function, although no cyclophilins were found during immunosolation of the TGN (Drakakaki et al., 2012).

The prenylated RAB acceptor B2 (PRA1.B2, AT2G40380) is found in both Golgi proteomes (FFE and LOPIT) but not the TGN, implying involvement with cisternal-specific interactions and vesicle docking. Examining proteins present uniquely in the TGN, besides those involved in trafficking such as the RAB GTases, soluble N-ethylmalimide-sensitive factor attachment protein receptors (SNARE; Blatt et al., 1999; Surpin and Raikhel, 2004), transport protein particle (TRAPP) components (Barrowman et al., 2010) or present as cargo, e.g., specific cellulose synthase A (CESA) subunits (Pavez et al., 2006), one endomembrane protein/transmembrane 9 protein (EMP/TMN9) and two S-adenosyl-l-methionine-dependent methyltransferases appear to stand out. Most EMP/TMN9 proteins are found in the Golgi cisternae: 11 members from a total of 12 were identified in FFE-purified samples (Parsons et al., 2012a) and 10 during LOPIT studies. EMP/TMN9 proteins interact with COP1 and COP10 proteins and membrane proteins destined for post-Golgi locations but are only recently studied in plants (Gao et al., 2012).

The employing the relational capabilities of the SUBA database, it is possible to compute a size estimate of the Golgi/TGN proteome based on each algorithms performance. The overall performance of each prediction program can vary considerably with regard to the total predicted "Golgi" proteins in Arabidopsis (contrast Adaboost, 6th Golgi and PProwler, 8885 Golgi) and positive prediction rate of the experimental proteome (contrast Adaboost <1% and PProwler >50%). However, after calculating false positive and false negative rates for each program, the final predicted Golgi proteomes are remarkably similar. Based on this analysis, the Arabidopsis Golgi/TGN proteome is estimated to be 2239 ± 465, employing the average of the predicted proteomes of these 14 subcellular prediction programs.

**USING THE PROTEOME: WHAT ARE THE ROLES OF UNCHARACTERIZED PROTEIN FAMILIES?**

A number of large gene families have been identified by both the FFE and LOPIT studies (Nikolovski et al., 2012; Parsons et al., 2012a). The quantitative mass spectrometry performed when applying LOPIT (Nikolovski et al., 2012) and spectral counts from FFE isolates (Parsons et al., 2012a), combined with localization data (Heazlewood et al., 2007), provide an important starting guide as to which members of these large families are major components and should be initially investigated in future studies.

The cyclophilin-like peptidyl-prolyl cis-trans isomerase family is consistently represented in the Golgi proteomes. These are known to catalyze conversion of cis to trans conformation of peptide bonds preceding prolyl residues in newly synthesized peptides (Chou and Gasser, 1997). In plants, they are classically associated with the thylakoid lumen where they are thought to help protein folding and assembly of photosystem complexes although their exact role is not clear (Ingelsson et al., 2009). The cyclophilins found by both FFE and LOPIT approaches (Nikolovski et al., 2012; Parsons et al., 2012a) localize either exclusively to the Golgi or are dually localized to the Golgi and plasma membrane (Dunkley et al., 2006; Benschop et al., 2007; Marmagne et al., 2007; Parsons et al., 2012a), implying a secretory-specific function, although no cyclophilins were found during immunosolation of the TGN (Drakakaki et al., 2012).

The prenylated RAB acceptor B2 (PRA1.B2, AT2G40380) is found in both Golgi proteomes (FFE and LOPIT) but not the TGN, implying involvement with cisternal-specific interactions and vesicle docking. Examining proteins present uniquely in the TGN, besides those involved in trafficking such as the RAB GTases, soluble N-ethylmalimide-sensitive factor attachment protein receptors (SNARE; Blatt et al., 1999; Surpin and Raikhel, 2004), transport protein particle (TRAPP) components (Barrowman et al., 2010) or present as cargo, e.g., specific cellulose synthase A (CESA) subunits (Pavez et al., 2006), one endomembrane protein/transmembrane 9 protein (EMP/TMN9) and two S-adenosyl-l-methionine-dependent methyltransferases appear to stand out. Most EMP/TMN9 proteins are found in the Golgi cisternae: 11 members from a total of 12 were identified in FFE-purified samples (Parsons et al., 2012a) and 10 during LOPIT studies. EMP/TMN9 proteins interact with COP1 and COP10 proteins and membrane proteins destined for post-Golgi locations but are only recently studied in plants (Gao et al., 2012).
| Predictor          | Predicted Golgi Arabidopsis | Expt. any location (575) | Expt. in Golgi (575%) | Expt. non-Golgi | FPR Golgi prediction | Ext. correct predictions | FNR Golgi prediction | Predicted Golgi | Non-predictable expt. Golgi | Reference                  |
|--------------------|-----------------------------|--------------------------|-----------------------|----------------|----------------------|------------------------|----------------------|----------------|----------------------------|---------------------------|
| AdaBoost           | 66                          | 13                       | 2                     | 11             | 0.85                | 10                     | 100                   | 2919           | 2909                       | Niu et al. (2008)          |
| ScuCG2             | 515                         | 100                      | 104                   | 997            | 0.91                | 489                    | 6.2                  | 2704           | 2215                       | Pierleoni et al. (2006)    |
| EpiLoc             | 210                         | 121                      | 30                    | 91             | 0.75                | 72                     | 0.95                 | 1371           | 1306                       | Brady and Shatkay (2008)   |
| IPSORT             | 588                         | 1439                     | 177                   | 1362           | 0.88                | 700                    | 0.69                 | 2273           | 1573                       | Bannai et al. (2002)       |
| MultiLoc2          | 916                         | 299                      | 114                   | 185            | 0.62                | 349                    | 0.80                 | 1702           | 1412                       | Blum et al. (2008)         |
| PlantinPred        | 1074                        | 310                      | 100                   | 230            | 0.30                | 325                    | 0.83                 | 1871           | 1546                       | Cheu and Shian (2010)      |
| PredAnno           | 596.1                       | 1438                     | 186                   | 1252           | 0.87                | 767                    | 0.68                 | 2410           | 1563                       | Small et al. (2004)        |
| PrediGlu           | 8760                        | 2012                     | 219                   | 1823           | 0.98                | 1092                   | 0.55                 | 2435           | 1333                       | Blum et al. (2008)         |
| PProver            | 8895                        | 21.21                    | 296                   | 1825           | 0.96                | 1240                   | 0.49                 | 2409           | 1169                       | Bodan and Hawkins (2009)   |
| SLPPA              | 7318                        | 1733                     | 142                   | 1591           | 0.92                | 634                    | 0.75                 | 2667           | 1933                       | Tamura and Akutsu (2007)   |
| SLPLocal           | 7408                        | 1599                     | 167                   | 1432           | 0.90                | 773                    | 0.71                 | 2683           | 1890                       | Matuda et al. (2009)       |
| TargetP            | 682                         | 15.97                    | 200                   | 1314           | 0.87                | 825                    | 0.65                 | 2372           | 1547                       | Emanuelli et al. (2000)    |
| WolfPSORT          | 112                         | 45                       | 11                    | 34             | 0.76                | 27                     | 0.98                 | 1431           | 1404                       | Horan et al. (2007)        |
| YLoc               | 1235                        | 318                      | 105                   | 213            | 0.67                | 408                    | 0.82                 | 2333           | 1625                       | Blommeister et al. (2008)  |

Predicted Golgi Arabidopsis: All proteins predicted to be Golgi in Arabidopsis (TAIR10).
Expt. any location (575): Predicted Golgi and experimentally determined to be in any location by MS or FP (SUBA) or Golgi (575).
Expt. in Golgi (575): Predicted Golgi but experimentally found to be in Golgi (575).
Expt. non-Golgi: Predicted Golgi but experimentally found to be non-Golgi.
FPR Golgi prediction: False positive rate for Golgi prediction (Expt. non-Golgi)/(Expt. in Golgi).
Ext. correct predictions: Estimation of correct predictions from total Golgi predictions in Arabidopsis (Expt. any location) – (Expt. any location) × (FPR Golgi prediction).
FNR Golgi prediction: False negative rate for Golgi prediction (1 – Expt. in Golgi)/(575).
Predicted Golgi: The predicted size of the proteome based on validated performance for each predictor program (Expt. any location) × (1 – FPR)(1 – FNR).
Non-predictable expt. Golgi: The size of the unpredictable Golgi proteome (Predicted Golgi) – (Ext. correct predictions).
Parsons et al. (2012a) but not LOPIT. No clear pattern, e.g., protein abundance, exists between the proteins observed in either study; most probably differences arise from variations in methodologies, highlighting the value of multi-faceted approaches to proteomic characterization of the Golgi.

WHAT IS MISSING FROM THE EXPERIMENTAL GOLGI PROTEOME?

Specific questions concerning what has not been identified so far are obviously difficult to answer but they can be addressed in part by examining what sorts of protein have been localized by fluorescent tagging but not identified by subcellular proteomic techniques. Fluorescent localization of proteins is generally motivated by interest in a specific protein and so is more likely to represent low-abundant polypeptides. It therefore provides an initial guide to the completeness of subcellular proteomic approaches.

Notably absent from proteomic surveys, but localized to the Golgi stack by fluorescent tagging are the Golgins and GRIP domain proteins (Latijnhouwers et al., 2007). Several glycosyltransferases such as cellulose synthase-like D5 (CSDL5; Bernal et al., 2007), rhamnogalacturonan II xyloolxyltransferase (RGXT) 1 and 2 (Egelund et al., 2006), irregular xylem 9 (IRX9; Pena et al., 2007), reversibly glycosylated polypeptide (RGP)1–4 (Drakakaki et al., 2006; Rautengarten et al., 2011), galacturonic acid transferase like (GATL) members from the GT8 family and a number of small GTPases are also either absent or poorly represented. Common methodological steps between these technically very different proteomes may in part explain these absences. Both the FFE and LOPIT approaches (Nikolovski et al., 2012; Parsons et al., 2012a) used cell suspension cultures whilst the immunosol- lation approach (Drakakaki et al., 2012) used 14-day-old liquid grown plantlets as the starting tissue, meaning that all proteomes were based on primary cell wall-rich tissue. This may explain the absence of CSDL5 and IRX9, which are both implicated in secondary cell wall biosynthesis and localized to the Golgi stack (Bernal et al., 2007; Lee et al., 2007). RGXT1 and 2 may have been also have been missed because of tissue-specific or low expression (Egelund et al., 2006). Members of the GATL clade, although localized to the Golgi stack (Kong et al., 2011), are absent from all Golgi proteomes, which could point toward some specific spatial or temporal function of these glycosyltransferases. Golgins are Golgi matrix proteins with coiled coil domains that typically locate to the cis- and trans-extremities of the Golgi stack and cisternal peripheries. They are involved in regulation of stack architecture and tethering events during trafficking (Osterrieder, 2012). Their location to cis- or trans-extremities of the Golgi stack may have precluded detection (Nikolovski et al., 2012; Parsons et al., 2012a). Peripheral golgins and those with GRIP domains which localize to the TGN, have no predicted transmembrane architecture and tethering events during trafficking (Osterrieder, 2012). It is conceivable that with a few more post-Golgi compartments characterized, many of the endomembrane proteins currently assigned to multiple locations (Headwood et al., 2007) could be reassigned and more light shed on the various protein cycling routes through the secretory pathway. This could be reasonably achieved in a number of ways. For the smaller compartments such as endosomal compartments, the immunosol- lation approach (Drakakaki et al., 2012) would hold the most promise as a number of syntaxin proteins known to associate with this compartment have been identified (Sanderfoot and Raikhel, 1999). Such an approach may not be appropriate for isolating individual cisternae from the main stack as trafficked proteins destined for later cisternae and TGN may also be detected by antibodies, whilst stack architecture could prove too complex for such an approach. Several fractions containing a high proportion of known Golgi proteins were not included in the FFE proteome owing to slightly higher level of contaminants. The number of fractions in which over 25% of proteins had been localized to the Golgi by LOPIT
Although one of the most technically challenging organelles to study, the Golgi complex has a unique role in the organelle network. It acts as a regenerative system for the cell, serving as a sorting and packaging center for proteins and lipids. The Golgi complex is responsible for the modification and sorting of proteins and lipids, which are then transported to various cellular compartments, including the plasma membrane. This process involves the formation of specialized transport vesicles that carry these molecules through the endomembrane system. The Golgi complex is the final sorting station for proteins and lipids, which are then released to their respective destinations.

CONCLUDING REMARKS

Although the Golgi complex is a relatively small organelle, its role in cellular biology is significant. It is responsible for the modification and sorting of proteins and lipids, which are then transported to various cellular compartments. The Golgi complex is a highly dynamic organelle, and its function is essential for the normal functioning of the cell. It is a versatile organelle that can adapt to different cellular environments and can respond to changes in the cell's physiological state. The Golgi complex is a key player in the cellular sorting and packaging process, and its function is essential for the normal functioning of the cell.

ACKNOWLEDGMENTS

This work was supported by the Joint BioEnergy Institute, which is supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The work conducted by Drakakaki et al. (2012) was funded by a grant from DOE (DEFG03-02ER15295) and UC Davis startup funds.
A novel representation of protein sequences for prediction of subcellular location using support vector machines. Protein Sci. 14, 2804–2813.

Mouras, S., Vorn, J. P., Sampa, H., Usda, N., Tobi, H., and Skirren, T. (2015). A novel representation of protein sequences for prediction of subcellular location using support vector machines. Protein Sci. 14, 2804–2813.

Moros, D. J., and Molleman, H. H. (1994). Isolation of Golgi apparatus from plant cells. J. Cell Biol. 123, 295–303.

Moros, D. J., and Molleman, H. H. (2009). The Golgi apparatus: The First 25 years. Trends Plant Sci. 14, 434–440.

Mouillot, C., Rabet, M. C., Carvalho, C., Hlad, C., Effert, D., Hennay, K., et al. (2007). Homogalacturonan synthesis in Arabidopsis thaliana requires a Golgi-localized protein with a putative methylntransferase domain. Plant J. 50, 605–614.

Nakamura, Y., Nishimura, Y., and Hotta, T. (2000). Utilization of amino acid composition for rapidly screening proteomes for their spatial arrangement along the secretory pathway. Plant Cell 20, 2219–2229.

Ontiveros, A. (2012). Talons of tethers and tentacles: golgins in plants. J. Membr. Biol. 247, 68–77.

Pendharkar, A. R., Somervaille, C. R., and Ehlerding, D. W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. Science 312, 1491–1495.

Parsons, H. T., Christiansen, K., Kusnetz, B., Carroll, A., Ito, J., Barth, T. S., et al. (2012a). Isolation and proteomic characterization of the Arabidopsis Golgi proteome. Plant Physiol. 159, 12–26.

Parsons, H. T., Greaves, N., and Heazlewood. This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science. Copyright © 2013 Parsons, Drakakaki and Heazlewood. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution, and reproduction in other forums, provided the original authors and source are credited and any change or alteration is properly noted concerning any third party graphics ex.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 November 2012; accepted: 12 December 2012; published: online 05 January 2013. Citation: Parsons HT, Drakakaki G and Heazlewood JL (2013) Proteomic dissection of the Arabidopsis Golgi and trans-Golgi network. Front. Plant Sci. 4:298. doi: 10.3389/fpls.2012.00298

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Parsons, Drakakaki and Heazlewood. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution, and reproduction in other forums, provided the original authors and source are credited and any change or alteration is properly noted concerning any third party graphics ex.
Author/s: Parsons, HT; Drakakaki, G; Heazlewood, JL

Title: Proteomic dissection of the Arabidopsis Golgi and trans-Golgi network

Date: 2013-01-03

Citation: Parsons, H. T., Drakakaki, G. & Heazlewood, J. L. (2013). Proteomic dissection of the Arabidopsis Golgi and trans-Golgi network. FRONTIERS IN PLANT SCIENCE, 3 (JAN), https://doi.org/10.3389/fpls.2012.00298.

Persistent Link: http://hdl.handle.net/11343/257726

File Description: Published version

License: CC BY