A genomic perspective on membrane compartment organization

Jason B. Bock, Hugo T. Matern, Andrew A. Peden & Richard H. Scheller

Howard Hughes Medical Institute, Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, California 94305-5426, USA

Now that whole genome sequences are available for many eukaryotic organisms from yeast to man, we can form broad hypotheses on the basis of the relative expansion of protein families. To investigate the molecular mechanisms responsible for the organization of membrane compartments, we identified members of the SNARE, coat complex, Rab and Sec1 protein families in four eukaryotic genomes. Of these families only the Rab family expanded from the unicellular yeast to the multicellular fly and worm. All families were expanded in humans, where we find 35 SNAREs, 60 Rabs and 53 coat complex subunits. In addition, we were able to resolve the SNARE class of proteins into four distinct subfamilies.

We compared families of proteins involved in vesicle trafficking from four organisms whose genome sequences are mostly complete: *Saccharomyces cerevisiae* (yeast)[2], *Drosophila melanogaster* (fly)[3,4], *Caenorhabditis elegans* (worm)[5,6] and *Homo sapiens* (human) (Table 1). Although the genomes are at various stages of completion, we feel that these results are an early view of the genome, not a premature one.

Vesicle-trafficking pathways

A hallmark of eukaryotic cells is the ability to segregate biochemical reactions within membrane-bound organelles, such as the endoplasmic reticulum, Golgi complex, endosomes and lysosomes. These compartments contain both resident proteins, which define each compartment according to its biochemical function, and proteins transiently passing through on the way to their final destinations. Cargo is loaded into a newly forming vesicle, which buds from a donor organelle and is transported to an acceptor organelle. Here it docks and fuses, delivering its contents (Fig. 1). These distinct steps are present in all vesicle-trafficking pathways, whether moving a protein from the endoplasmic reticulum to the Golgi for glycosylation or from an endosome to a lysosome for degradation. A central issue in cell biology has been to understand how a cell carries out the various steps in the life cycle of a transport vesicle, and how specificity is conferred to the system, so that, for example, an endoplasmic reticulum-derived vesicle delivers its cargo to the Golgi and not the lysosome[1].

The protein families that mediate vesicle trafficking are conserved through phylogeny from yeast to man, as well as throughout the cell from the endoplasmic reticulum to the plasma membrane (see Supplementary Information). Each vesicle-trafficking step involves a similar set of proteins drawn from these families. By having different members of these families localized to distinct membrane compartments, specificity is encoded in the machinery itself.

Genome analysis

The biogenesis of transport vesicles is initiated through the recruitment of large multi-subunit protein complexes termed coats[7]. These coats are central to both cargo selection and deformation of the lipid bilayer into a budding vesicle (Fig. 1, **1**). Each coat complex is recruited to a distinct membrane compartment within the cell, indicating that individual coats are involved in specific transport steps. A coat complex is formed by the interaction of several different subunits. For example, an AP-1 adapter complex, involved with transport from the trans-Golgi network to endosomes, consists of four subunits (γ, β1, μ1 and σ1). Individual subunits carry out discrete functions within the complex, effecting the localization, regulation and affinity of the complex for cargo molecules.

Our analysis reveals that the number of coat complexes has increased only slightly from yeast, fly, and worm to human (from six to seven); however, the number of individual coat subunits has expanded significantly in humans (Table 1). This indicates that, instead of evolving new coat complexes for specialized trafficking steps and to accommodate increased cargo complexity, mammals have used a modular system where specificity can be achieved through subunit exchange. For example, the adaptor medium chains μ1A and μ1B are very similar, but only AP-1 adaptor complexes containing μ1B can efficiently transport low-density lipoprotein (LDL) and transferrin receptors to the basolateral surface in polarized cells.
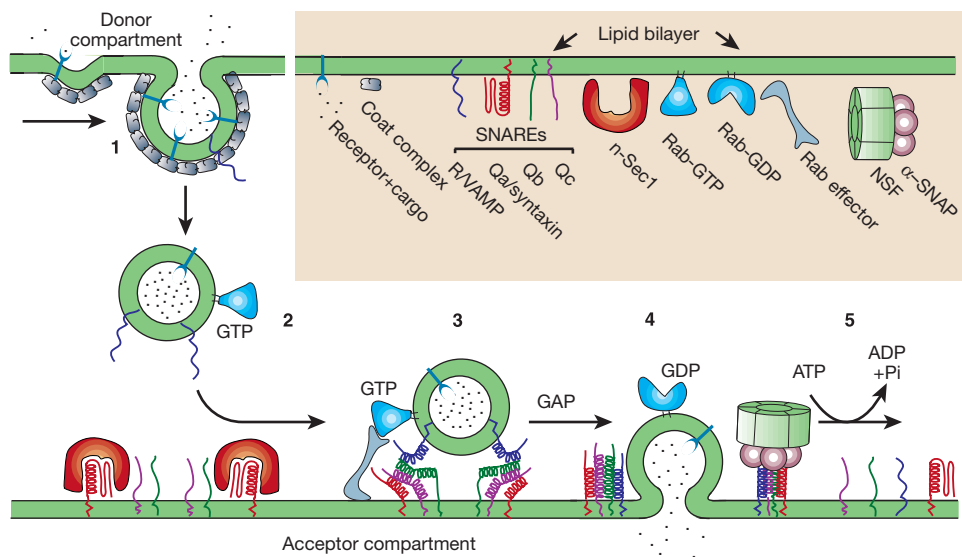
The next steps in the life cycle of a transport vesicle involve proteins of the Rab (Ypt in yeast) family. Rabs are a subfamily of the Ras superfamily of low-molecular-mass GTPases and cycle between a GTP, membrane-bound, active state and a GDP, cytosolic, inactive state. When bound to GTP, Rabs interact with a host of different proteins, loosely termed Rab effectors[8]. These Rab effectors perform diverse functions from vesicle budding to vesicle transport by way of the cytoskeleton, and vesicle docking at the target membrane (Fig. 1, **1–3**). Different Rabs are localized on distinct vesicles and organelles, and are positioned to demarcate particular membranes. Thus, through their localization and GTP state, Rabs can recruit and/or activate their various effectors at the correct time and to the correct place, providing an element of regulation to the vesicle-trafficking machinery.

The number of Rabs in the four eukaryotic genomes analysed scaled roughly with the total number of predicted genes. The human

Table 1 Number of members of vesicle-trafficking families

	S. cerevisiae	*C. elegans*	*D. melanogaster*	*H. sapiens*
Coat complexes (subunits)	6 (31)	6 (29)	6 (29)	7 (53)
Rabs	11	29	26	60
SNAREs	21	23	20	35
Qa-SNAREs/syntaxins	7	9	7	12
Qb-SNAREs/SNAP Ns	5	7	5	9
Qc-SNAREs/SNAP Cs	6	4	5	8
R-SNAREs/VAMPs	5	6	5	9
Sec1s	4	6	5	7
Total predicted genes[4]	6,241	18,242	13,601	30,000–50,000

For a more extensive analysis of these protein families, including accession numbers and searching methodology, see Supplementary Information. SNARE helical definitions were determined by protein profiling. Some proteins (SNAP-25) contain two SNARE-coil domains and thus are counted twice in the coil subdivisions.

Figure 1 The life cycle of a transport vesicle. **1**: Budding. A small Arf/Sar GTPase is recruited to the donor membrane by its cognate guanine nucleotide exchange factor (see Supplementary Information). Next, incoming coat components are bound and cargo proteins diffuse into the budding site where they are trapped by their interaction with coat proteins or cargo receptors. The membrane forms a coated bud, giving rise to a vesicle. A GTPase activating protein (GAP) that is part of the coat stimulates GTPase activity, which leads to uncoating and release of the GTPase and coat proteins into the cytosol, where they can recycle. **2**: Transport. Vesicles are transported from their budding site to the acceptor compartment through association with cytoskeletal elements and transport motors (not shown). **3**: Docking. Vesicles are targeted to sites of fusion in an event that involves a GTP-bound Rab protein, elongated coiled-coil tethering proteins and other tethering factors (not shown). After the docking step, an R-SNARE (VAMP) assembles with a Qa-SNARE (syntaxin), a Qb-SNARE (SNAP N) and a Qc-SNARE (SNAP C) to form a parallel, four-helical bundle. Interactions of Sec1s with Qa-SNAREs are probably critical in forming the core fusion complex. **4**: Fusion. After nucleation of the ternary core fusion complex, further zippering of the parallel helices brings the vesicle and acceptor membrane close together, perhaps driving the fusion reaction. **5**: Disassembly. After fusion, Rab-GDP is released from the membrane by GDP-dissociation inhibitor (GDI) and recruited back to the donor membrane. The SNARE complex is disassembled by the ATPase N-ethylmalemide-sensitive factor (NSF) and α-SNAP so that SNAREs can be recycled for further rounds of transport.

genome contains 60 Rabs, including 31 that were identified in this study (23 have known mammalian orthologues). As Rabs perform most of their functions by interacting with their effectors, until we have a more complete catalogue of those effectors and GDP/GTP exchange factors, we will not have a firm grasp on all of their functions. It has been difficult to identify Rab effectors because they share little, if any, sequence homology. This is a case where traditional genomics is of little help, but where the coming era of structural genomics may have an impact[9]. There is evidence that Rabs have both an essential role and a more subtle, regulatory role in vesicle trafficking. In yeast, Ypt1 (Rab1) is essential for trafficking from the endoplasmic reticulum to the Golgi apparatus. In contrast, Rab3A, present on synaptic vesicles, appears to have a regulatory role. Rab3A knockout mice are viable, but certain synapses in their brains fail to undergo a type of plasticity called long-term potentiation. Thus, whereas some Rabs appear to be essential parts of core vesicle-trafficking machinery, others appear to act more as regulators of the process. Perhaps an expansion in the regulatory role of Rabs was necessary to coordinate cell physiology with membrane trafficking in complex, multicellular organisms.

The final step in vesicle trafficking is the fusion of a vesicle with its target membrane, believed to be mediated by a family of proteins termed SNAREs[1]. SNAREs are integral membrane proteins present on both vesicle and target membranes, and can form very stable complexes. The formation of a SNARE complex pulls the vesicle and target membrane together and may provide the energy to drive fusion of the lipid bilayers. It is proposed that a unique SNARE complex, consisting of four helices, is formed at each fusion site. For example, the fusion of synaptic vesicles with the presynaptic nerve terminal is mediated by the formation of a complex comprising one helix each from syntaxin 1A and VAMP-2 and two helices from

SNAP-25. These four helices are prototypes that can be used to define subclasses of SNAREs on the basis of protein profiling (see Table 1 and Supplementary Information). Different members of the SNARE families are localized to distinct membrane compartments throughout eukaryotic cells and thus are positioned to form specific complexes and to enhance the fidelity of vesicle trafficking at the final step.

The SNARE family has remained mostly unchanged in yeast, flies and worms, but has increased significantly in humans to 35 members. This leads to the conclusion that multicellular organisms do not have an inherently more complex secretory pathway, and that a set of core SNAREs is sufficient to mediate most intracellular vesicle fusion events. Mammals appear to have taken multicellular specialization to another level, with differential expression of more SNARE proteins. For example, syntaxin 17 is abundantly expressed in steroidogenic tissues, and is localized to the smooth endoplasmic reticulum. Neurotransmission is a specialized form of exocytosis mediated by syntaxins 1A and B, which are greatly enriched in neurons. So, although multicellular organisms clearly can function with a core set of SNAREs, use of additional SNAREs may result in further tissue-specific specialization of membrane trafficking.

A protein family that interacts directly with the syntaxin subfamily of SNAREs is classified by the prototypical member Sec1 (also known as mUNC-18). Sec1s are cytosolic proteins, which are peripherally associated with membranes in part through their interactions with syntaxins. Sec1 is essential for vesicle trafficking; a knockout of *SEC1* in yeast abolishes exocytosis. The ratio of Sec1s to syntaxins in the genomes analysed remains roughly constant, supporting the idea that Sec1s act primarily through syntaxins. They are probably chaperones, putting syntaxins into conformations that are required for interactions with other SNAREs. As with

the SNAREs, the increase in Sec1 genes can be attributed to specialization of trafficking events. For example, the Sec1 family member mUNC-18c interacts with syntaxin 4 and is involved in the translocation of glucose transporters to the cell surface in response to insulin.

Conclusions

Analysis of the extent of four families of proteins involved in vesicle trafficking leads to several conclusions. The evolutionary leap from single to multicellular organisms did not require an increase in the core machinery (coats, SNAREs, Sec1s) underlying the transport vesicle life cycle. The Rab family did expand significantly from yeast to worms and flies, implying that multicellular organisms exert more regulation over vesicle-trafficking pathways. The evolutionary jump to mammals saw a significant increase in all of the families examined. This implies that mammals orchestrate the complexities of multicellular physiology through not only more finely tuned regulation, but also tissue-specific specialization of the core trafficking machinery. This is perhaps most clearly illustrated in the brain, where memories are at least partially encoded through regulation of specialized membrane trafficking proteins most abundantly expressed in neurons. □

1. Lin, R. C. & Scheller, R. H. Mechanisms of synaptic vesicle exocytosis. *Annu. Rev. Cell Dev. Biol.* **16,** 19–49 (2000).
2. Goffeau, A. *et al.* The Yeast Genome Directory. *Nature* **387,** (suppl.), 1–105 (1997).
3. Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster. Science* **287,** 2185–2195 (2000).
4. Rubin, G. M. *et al.* Comparative genomics of the eukaryotes. *Science* **287,** 2204–2215 (2000).
5. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology [published errata appear in *Science* **283,** 35 (1999), *Science* **283,** 2103 (1999) and *Science* **285,** 1493 (1999)]. *Science* **282,** 2012–2018 (1998).
6. Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J. & Spieth, J. Wormbase: network access to the genome and biology of *C. elegans. Nucleic Acids Res.* (in the press).
7. Hirst, J. & Robinson, M. S. Clathrin and adaptors. *Biochim. Biophys. Acta* **1404,** 173–193 (1998).
8. Novick, P. & Zerial, M. The diversity of Rab proteins in vesicle transport. *Curr. Opin. Cell Biol.* **9,** 496–504 (1997).
9. Skolnick, J., Fetrow, J. S. & Kolinski, A. Structural genomics and its importance for gene function analysis. *Nature Biotechnol.* **18,** 283–287 (2000).