Molecular & Biochemical Parasitology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Resolving the homology—function relationship through comparative genomics of membrane-trafficking machinery and parasite cell biology

Christen M. Klinger[a], Inmaculada Ramirez-Macias[a], Emily K. Herman[a], Aaron P. Turkewitz[b], Mark C. Field[c], Joel B. Dacks[a,*]

[a] Department of Cell Biology, University of Alberta, Edmonton, Alberta, Canada
[b] Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL, USA
[c] School of Life Sciences, University of Dundee, Dundee, UK

ARTICLE INFO

Article history:
Received 19 April 2016
Received in revised form 12 July 2016
Accepted 16 July 2016
Available online xxx

Keywords:
Membrane-trafficking
Protist
Parasite
Genomics
Functional homology
Endomembrane

ABSTRACT

With advances in DNA sequencing technology, it is increasingly common and tractable to informatically look for genes of interest in the genomic databases of parasitic organisms and infer cellular states. Assignment of a putative gene function based on homology to functionally characterized genes in other organisms, though powerful, relies on the implicit assumption of functional homology, *i.e.* that orthology indicates conserved function. Eukaryotes reveal a dazzling array of cellular features and structural organization, suggesting a concomitant diversity in their underlying molecular machinery. Significantly, examples of novel functions for pre-existing or new paralogues are not uncommon. Do these examples undermine the basic assumption of functional homology, especially in parasitic protists, which are often highly derived? Here we examine the extent to which functional homology exists between organisms spanning the eukaryotic lineage. By comparing membrane trafficking proteins between parasitic protists and traditional model organisms, where direct functional evidence is available, we find that function is indeed largely conserved between orthologues, albeit with significant adaptation arising from the unique biological features within each lineage.

1. Introduction

Genomics, the sequencing and analysis of genomes has empowered tremendous advances. Possessing a genome sequence for an organism, particularly one difficult to culture or genetically manipulate, allows the prediction of cellular organization, metabolism, gene expression mechanisms, and organellar complement, through *in silico* analysis of the corresponding predicted proteome.

This is essentially a comparative analysis, which at its heart relies on robust evidence of function in one or more organisms. Comparative genomics allows reconstruction of pan-eukaryotic complements of cellular components, including the cytoskeleton, nuclear transport, metabolism, and mitochondrion ([1], *inter alia*), providing evidence for the general or core aspects of cellular systems and which aspects are lineage-specific. This evidence is an important basis for understanding evolutionary mechanisms behind the emergence of cellular complexity. Furthermore, the acceleration in understanding gained by the annotation of thousands of genes is invaluable, by producing initial hypotheses for expected interactions, pathways, and organellar roles that can be tested.

Inherent in comparative genomic studies is the assumption of functional homology, *i.e.* that orthologous genes retain equivalent function. Orthology is the relationship between two genes in distinct taxa that are directly related by vertical descent [2], and which may be considered as the "same gene"; the expectation is that such gene pairs retain equivalent properties and roles within the cell [3]. This assumption has been generally regarded as safe, based on a model of conservation of function rather than the widespread gain of novel functions or neofunctionalization and based on experimental validation of enzymes assayed heterologously or *in vitro*, where 'function' can be relatively readily defined. However, much of our understanding of eukaryotic cell biology is based on evidence

Abbreviations: AP, adaptor protein; ESCRT, endosomal sorting complexes required for transport; SNARE, soluble *N*-ethyl-maleimide-sensitive factor attachment protein receptors; Rab, ras from brain; Vps, vacuolar protein sorting.

* Corresponding author at: Department of Cell Biology, University of Alberta, 5-31 Medical Science Building, Edmonton, Alberta, T6G 2H7, Canada.
E-mail address: dacks@ualberta.ca (J.B. Dacks).

http://dx.doi.org/10.1016/j.molbiopara.2016.07.003
0166-6851/

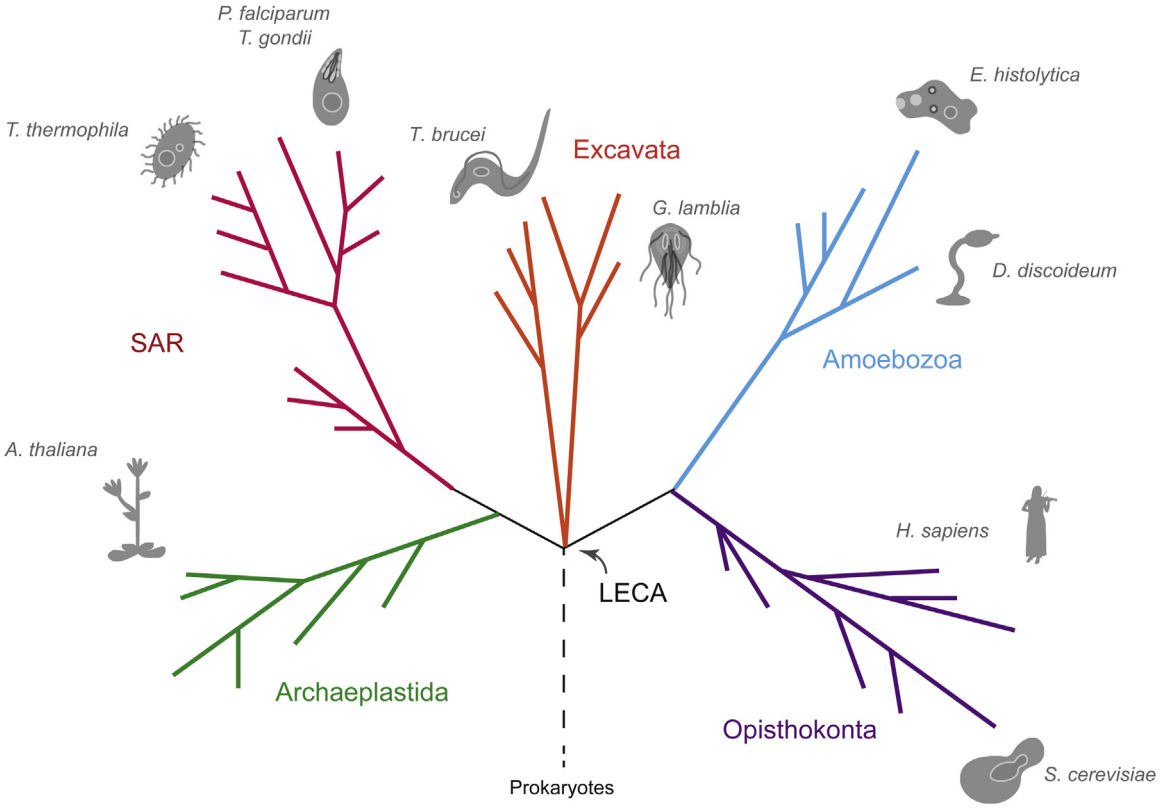

Fig. 1. Model Organisms Across Eukaryotes. This figure demonstrates the distribution of model organisms across eukaryotic diversity. Colour-coded branches and corresponding labels denote eukaryotic Supergroups, with the branching order roughly corresponding to the organization of known diversity within each group. Model organisms are represented by greyscale illustrations and corresponding labels in italics. The position of the Last Eukaryotic Common Ancestor (LECA) is indicated. Though additional model organisms exist for each of these groups, they are excluded from this figure for simplicity.

from a small sample of true eukaryotic diversity and frequently from a restricted region of the eukaryotic tree. Given this sampling bias, to what extent can 'function' be reliably predicted across eukaryotic diversity based on sequence similarity alone?

Testing the assumption of functional homology requires experimental evidence from organisms across a full taxonomic range of eukaryotes, and fortunately there are now tractable organisms from each of the major eukaryotic divisions or Supergroups (Fig. 1). Here we have chosen a subset of non-metazoan organisms and assessed comparative data available for genes of the membrane trafficking system, a crucial cellular system underpinning pathogenic mechanisms in many parasitic protists, and which has been well studied. We not only assess the validity of the core assumption of functional homology in comparative studies of membrane trafficking genes, but also begin to identify the manner in which the endomembrane system is modified in individual parasitic lineages and which speaks directly to mechanisms of disease and the origins of parasitism.

1.1. The membrane-trafficking system: a modern molecular view

Membrane trafficking is the process by which proteins and other macromolecules are distributed throughout organelles of the endomembrane system, and released into, or internalized from, the extracellular environment. Trafficking is vital for metabolism, signaling, and interacting with the external environment. Transport vesicles act to transfer cargo molecules between the organelles of the endomembrane system, which possess discrete morphology, localization, and functions [4].

Anterograde trafficking involves movement from the endoplasmic reticulum (ER) through the Golgi complex, the *trans*-Golgi network (TGN), and on to the plasma membrane [5], whilst endocytosis begins at the plasma membrane where cargo is sorted by endosomes before recycling or targeting to acidic terminal organelles. During endocytosis organelles acidify, may acquire intralumenal vesicles (present in multi-vesicular bodies or MVBs), and modify their compositions [6]. In all trafficking pathways, retrograde transport steps recycle selected components back to previous organelles for use in future rounds of trafficking (Fig. 2).

Specialized protein complexes controlling vesicle budding, tethering, and fusion, many of which are large paralagous families, regulate transport. Arf/Sar family small GTPases and their regulators, cargo adaptors, and coat protein complexes are involved in vesicle formation/fission. Rab GTPases are involved in vesicle targeting, whilst coiled coil SNARE proteins are central to vesicle fusion [4]. Importantly, members of these multiple families act at discrete locations or trafficking pathways; the specificity of trafficking is in part encoded in the combinatorial interactions of these various players [7]. For example, COPII-coated vesicles mediate anterograde transport from the ER to the Golgi, while the corresponding retrograde transport step requires COPI vesicle formation [8]; clathrin-coated vesicles mediate multiple post-Golgi transport routes [9].

Our view of membrane trafficking is dominated by studies in animal and yeast cells. However, membrane trafficking is a central process underpinning growth, cell surface presentation and secretion. Thus it is critical to pathogenic mechanisms of many parasitic protists, for example by mediating host cell invasion [10] and immune system evasion [11]. It is therefore reasonable to ask what complement of membrane trafficking proteins is present across the broad diversity of eukaryotes and what we can infer about both evolution of the membrane trafficking system and the conserved set

Please cite this article in press as: C.M. Klinger, et al., Resolving the homology—function relationship through comparative genomics of membrane-trafficking machinery and parasite cell biology, Mol Biochem Parasitol (2016), http://dx.doi.org/10.1016/j.molbiopara.2016.07.003