12. Beukeboom, L.W., and Perrin, N. (2014). The Evolution of Sex Determination (Oxford: Oxford University Press).

13. Graves, J.A.M. (2013). How to evolve new sex determining genes. Dev. Dyn. *242*, 354–359.

14. Charlesworth, D. (2015). Plant contributions to our understanding of sex chromosome evolution. New Phytol. *208*, 52–65.

15. Charlesworth, D. (2013). Plant sex chromosome evolution. J. Exp. Bot. *64*, 405–420.

16. Marais, G.A.B., Nicolas, M., Bergero, R., Chambrier, P., Kejnovsky, E., Moneger, F., Hobza, R., Widmer, A., and Charlesworth, D. (2008). Evidence for degeneration of the Y chromosome in the dioecious plant *Silene latifolia*. Curr. Biol. *18*, 545–549.

17. Bergero, R., Qui, S., and Charlesworth, D. (2015). Gene loss from a plant sex chromosome system. Curr. Biol. *25*, 1234–1240.

18. Papadopulos, A.S.T., Chester, M., Ridout, K., and Filatov, D.A. (2015). Rapid Y degeneration and dosage compensation in plant sex chromosomes. Proc. Natl. Acad. Sci. USA *112*, 13021–13026.

Evolution: Causality and the Origin of Parasitism

Jan Janouskovec[1] and Patrick J. Keeling[2]
[1]Department of Biology, San Diego State University, San Diego, CA, USA
[2]Department of Botany, University of British Columbia, Vancouver, BC, Canada
Correspondence: janjan.cz@gmail.com (J.J.), pkeeling@mail.ubc.ca (P.J.K.)
http://dx.doi.org/10.1016/j.cub.2015.12.057

The first comparison of parasitic trypanosomatids to their free-living relatives reveals that some characteristics once linked to parasitism actually predate it. Parallel comparisons of other parasites suggest we need to rethink the drivers and consequences of the parasitic lifestyle.

Evolution works in mysterious ways, and one problem that biologists must confront is the difficulty in establishing causality. Understanding a biological transition means understanding not only what events took place, but also their order, and which events precipitated others. But we often predict order and causality by pure intuition; we observe characteristics in some organism or lineage that operate under a current set of conditions and then try to imagine a progression of events that gave rise to these characteristics. Often, however, we put the causal cart before the horse because we lack critical information, typically relating to the nature of the ancestral and intermediate states in this transition. Our understanding of the evolution of parasites and pathogens illustrates this problem all too well. While we have established concepts like the Bradford Hill criteria and Koch's postulates to at least guide our thinking about disease causation [1], understanding causal factors in the emergence of pathogens at evolutionary time scales has been more *ad hoc*. One reason for this is because we focus on the final state of this transition, the pathogen, and lack almost any information on the initial state, represented today by the pathogen's free-living relatives. In a recent paper in *Current Biology*, Jackson *et al.* [2] approach the origin of trypanosomatids, an important group of kinetoplastid parasites including *Trypanosoma* and *Leishmania*, by characterizing the genome of a free-living relative, the bacterial grazer *Bodo saltans*. By looking from the outside in, Jackson *et al.* [2] describe a number of noteworthy patterns about the kinetoplastid transition to parasitism that upend more than a few ideas about what led the kinetoplastids to become parasites.

Parasites abound with specialized characteristics associated with disease, and these are commonly interpreted as adaptations to the parasitic lifestyle. This may be true, but it may also be a reflection of our tendency to associate current specialization with the immediate biological context, and not with the context in which it first arose.

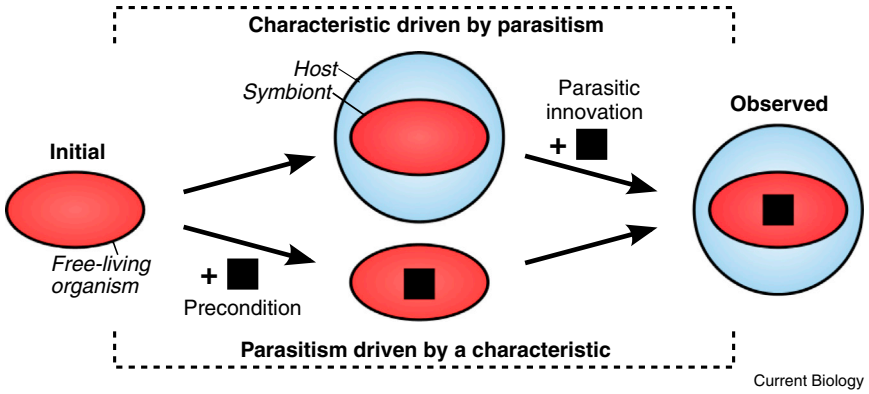

Figure 1. Models for the transition of free-living organisms to parasites, and their character states.
Two alternative ways to interpret the characteristics of parasites and how they relate to the origin of parasitism. On the top, a characteristic evolves as an adaptation to the existence of a parasitic interaction. On the bottom, the characteristic evolves in a free-living organism, but its existence makes the transition to parasitic life more likely, so it preconditions the lineage to a parasitic transition.

Table 1. Examples of apicomplexan and trypanosomatid characteristics reconstructed in recent studies.

	Pre-parasitic candidate preconditions*	Ancestral parasitic candidate early parasitic innovations**	Lineage-specific candidate late parasitic innovations
Apicomplexans see [5,6] for protein names and acronyms	**Gains**: Pseudoconoid, rhoptry and microneme-like organelles; schizogony-like division and ISPs; resting or reproductory cysts and oocyst wall proteins; many unique horizontally-acquired proteins (e.g., PREX, FECH, R2e2); RAP transcriptional regulation pattern; GAPMs; alveolin and subtilisin diversification	**Gains**: Gliding motility: MyoA, myosin class XIV, alveolar fusion; conoid closure; four rhoptry proteins **Losses**: Biosynthesis of purines, sterol, cofactors, aromatic amino acids; glyoxalate cycle; five IFT and several endomembrane components	**Gains**: Most proteins implicated in interaction with the host: e.g., ROP, RON, MIC, GRA families; diversification of apiAP2 domains **Losses:** Asymmetric: more in blood specialists (*Plasmodium*, *Babesia*) less in generalists (*Toxoplasma*)
Trypanosomatids see [2] for protein names and acronyms	**Gains**: Bodonin family of transmembrane glycoproteins; polycistronic transcription unit-like gene arrangement with conserved strand-switch regions; spliced leader trans-splicing **Losses:** Biosynthesis of purines; glutathione redox system; multiple endomembrane trafficking components (e.g., SNARE)	**Gains**: Diversification of amino acid and nucleoside transporters, amastins, cathepsin-L, calpains **Losses**: Catabolism: glucosidases, lipases, proteases; ion channels; purine biosynthesis; glutathione redox system; ABC transporters	**Gains**: Most proteins implicated in interaction with the host, e.g. cell surface VSG, PSA, DGF-1, mucin, trans-sialidase, and MSP families **Losses:** Asymmetric: cathepsins-B, transposons (VIPER, TATE)

Characteristics may be better understood as pre-parasitic (present in parasites and free-living relatives), ancestral parasitic (confined to parasites including their deepest lineage), and lineage-specific (confined to a subgroup of parasites). Pre-parasitic traits may also represent later exaptations that did not contribute to establishing their early dependency on the host, whereas ancestral parasitic traits may have contributed to the dependency on the host. Abbreviations: ISP, inner membrane complex (IMC) subcompartment protein; RAP, rhoptry associated protein; GAPM, glideosome-associated protein with multiple-membrane spans; IFT, Intraflagellar transport.

For example, a function we see as a 'derived' adaptation to parasitism may really have originated prior to parasitism; thus parasitism may either be irrelevant to the origin of that function, or the origin of parasitism may have even been facilitated by that function (Figure 1). Trypanosomatids certainly do not lack for strange, stand-out characteristics [3,4], and because these characteristics are seemingly 'unique' to these parasites, they are often presumed to be related to parasitism. However, this relationship must be reconsidered in light of the *Bodo* genome, which shows that many such characteristics evolved in free-living organisms, prior to the origin of parasitism [2]. Like trypanosomatids, *Bodo* largely lacks introns, and almost one tenth of its genes are co-linear to those in the parasites, a level unseen among other eukaryotes of comparable taxonomic distance. The explanation for these features lies in the shared presence of long polycistronic transcription units (PTUs), where many genes are transcribed unidirectionally in one go. This organization, which is very rare in eukaryotes, is itself promoted by the presence of another oddity, spliced leader trans-splicing, which facilitates processing of the polycistronic messages into single-gene transcripts. Also like trypanosomatids, *Bodo* lacks purine biosynthesis, the glutathione redox system, and several components of endomembrane trafficking. Indeed, *Bodo* does not contain any complete metabolic pathways that are absent from its parasitic relatives, suggesting that *Bodo* acquires metabolites from its prey much like the parasites do from their hosts. Perhaps most interestingly, the *Bodo* genome encodes a large family of predicted outer-membrane glycoproteins called bodonins, which are similar in structure to trypanosomatid surface proteins involved in host evasion (for example, the *Trypanosoma cruzi* Dispersed Gene Family-1 proteins and *Leishmania* Promastigote Surface Antigens). Overall, the *Bodo* genome shows that most of the characteristics that appeared to correlate with the parasitic life of trypanosomatids — and were accordingly viewed as potential adaptations to parasitism — are in reality much older than parasitism and evolved for completely different reasons.

These advances undoubtedly clarify our thinking on causality. However, to move even further towards asking whether any general principles exist and what they might be, we also need to examine multiple origins of parasitism from a similar perspective. Interestingly, many patterns observed by Jackson *et al.* [2] with regard to the origin of trypanosomatid parasitism appear to have parallels in another important lineage of parasites, the apicomplexans (which includes *Plasmodium*, *Toxoplasma*, and *Cryptosporidium)*. The closest cousins of apicomplexans are heterotrophic and photosynthetic species called the chrompodellids [5,6]. Like *Bodo*, chrompodellids collectively share a great number of characteristics with their parasitic kin. These include cellular structures like the pseudoconoid, rhoptry and microneme-like organelles, schizogony-like division, and thick-walled cysts, as well as molecular features such as gene families key to the parasites' success (such as the glideosome-associated protein with multiple-membrane spans and the oocyst wall proteins), protein domain architecture and transcriptional regulation

[6,7]. Before these features were recognized in the chrompodellids, their presence in the apicomplexans was interpreted as adaptations to parasitism.

These characteristics clearly did not originate within the context of parasitism. Some were probably not directly relevant to the origin of parasitism at all, even if they have since acquired some function related to infection and are best viewed as parasite exaptations (co-optations). Others might actually have played a different kind of role in the origin of parasitism. If a pre-parasitic function broadly corresponds to the function in a parasite, then the presence of that function could have been a precondition that made a shift to parasitism more likely (Figure 1 and Table 1). For example, organelles homologous to the apicomplexan infection machinery were used for feeding in the ancestors of apicomplexans and chrompodellids; so, although they evolved for feeding, their presence may have facilitated the emergence of parasitism. The presence of large gene families of surface coat proteins in both apicomplexans and kinetoplastids could similarly have had the same effect by facilitating their interactions with host immunity.

This is not to say that parasites have no distinguishing features. Examples include gliding motility, a true conoid (a variation in the open-sided feeding tube used by predators), and several rhoptry (specialized secretory organelle) proteins in apicomplexans, and diversification in amino acid transporters, nucleoside transporters, and amastins (surface glycoproteins) in trypanosomatids (Table 1). But what is striking is how few of these characteristics remain after even a first look at free-living relatives. This serves to focus attention on other kinds of change and other time scales that may be more significant. One example is the importance of lineage-specific innovations (Figure 1 and Table 1), which suggest that important aspects of parasitic life strategies were determined comparatively recently and are likely tied to host interactions. Another is the role of reduction and loss; these have not been overlooked, but they are difficult to interpret without knowing when they took place relative to the origin of parasitism. Metabolic losses are often somehow associated with parasitism, but in trypanosomatids this reduction is now seen to have preceded parasitism (evident in the similarities with *Bodo*) whereas in apicomplexans it did not (evident in the wider range of metabolic capabilities in chrompodellids). At the level of genome organization, the disconnect between parasitism and evolution of characteristics is even more pronounced. Jackson *et al.* [2] propose that the genome reduction in trypanosomatids is best viewed as a streamlining of gene families, which are compacted relative to *Bodo*, yet almost none being lost completely (most loss is associated with macromolecular degradation and membrane transport, consistent with the loss of phagotrophic feeding in the trypanosomatids). The authors suggest that this streamlining resulted from the transition to a more stable and physiologically narrower host environment, void of redundancy-maintaining selection. This is a reasonable assumption considering that similar patterns are found more broadly: in bacterial symbionts compared with free-living relatives [8,9], in specialist parasites compared with related generalists (for example, piroplasmid versus coccidean apicomplexans [10,11]), and even in free-living organisms, where specialists in narrow-niche environments streamline their genomes (see [12–14] for examples). Dramatic patterns of gene content reduction are not associated with the origin of parasitism per se, but rather are tied to broader characteristics of the environment. At the same time, evidence is mounting that other characteristics of a parasite's genomic architecture result from intrinsic characteristics of the genome itself (such as substitution rates and mutation biases), rather than from adopting a parasitic lifestyle [15,16].

Altogether, the emerging picture is that there is no simple genomic 'smoking gun' that betrays the evolutionary transition from a free-living lifestyle to a parasitic one, much less one that explains how or why this transition took place. Instead, the origin of parasitism is being steadily integrated into a larger story where ecological transitions are balanced with evolutionary change. Part of this process is re-thinking explanations so that the horse consistently precedes the cart, and this is partly enabled by emphasizing non-adaptive processes and preconditions. Lastly, although this discussion is focused on parasitism, we should point out that this is equally applicable to any symbiosis. Theories for the origin of chloroplasts, mitochondria, and eukaryotes themselves often build on the causality of adaptations [17,18] and, much like theories for the origin of parasites, this may overshadow the significance of preconditions in shaping organismal interactions.

REFERENCES

1. Hill, A.B. (1965). The environment and disease: association or causation? Proc. R. Soc. Med. *58*, 295–300.

2. Jackson, A.P., Otto, T.D., Aslett, M., Armstrong, S.D., Bringaud, F., Schlacht, A., Hartley, C., Sanders, M., Wastling, J.M., Dacks, J.B., *et al*. (2016). Kinetoplastid phylogenomics reveals the evolutionary innovations associated with the origins of parasitism. Curr. Biol. *26*, 161–172.

3. Lukes, J., Leander, B.S., and Keeling, P.J. (2009). Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates. Proc. Natl. Acad. Sci. USA *106*, 9963–9970.

4. Jackson, A.P. (2015). Genome evolution in trypanosomatid parasites. Parasitology *142*, S40–S56.

5. Janouškovec, J., Horák, A., Oborník, M., Lukeš, J., and Keeling, P.J. (2010). A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc. Natl. Acad. Sci. USA *107*, 10949–10954.

6. Janouškovec, J., Tikhonenkov, D.V., Burki, F., Howe, A.T., Kolísko, M., Mylnikov, A.P., and Keeling, P.J. (2015). Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. Proc. Natl. Acad. Sci. USA *112*, 10200–10207.

7. Woo, Y.H., Ansari, H., Otto, T.D., Klinger, C.M., Kolisko, M., Michálek, J., Saxena, A., Shanmugam, D., Tayyrov, A., Veluchamy, A., *et al*. (2015). Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. eLife *4*, e06974.

8. Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.-C., and Herrmann, R. (1996). Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res. *24*, 4420–4449.

9. Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Pontén, T., Alsmark, U.C., Podowski, R.M., Näslund, A.K., Eriksson, A.S., Winkler, H.H., and Kurland, C.G. (1998). The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature *396*, 133–140.

10. Reid, A.J., Blake, D.P., Ansari, H.R., Billington, K., Browne, H.P., Bryant, J., Dunn, M., Hung, S.S., Kawahara, F., Miranda-Saavedra, D., *et al*. (2014). Genomic analysis of the causative agents of coccidiosis in domestic chickens. Genome Res. *24*, 1676–1685.

11. Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., Carlton, J.M., Hall, N., Ren, Q., Paulsen, I.T., Pain, A., Berriman, M., *et al*. (2005). Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. Science *309*, 134–137.

12. Matsuzaki, M., Misumi, O., Shin-i, T., Maruyama, S., Takahara, M., Miyagishima, S.Y., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., *et al*. (2004). Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature *428*, 653–657.

13. Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., *et al*. (2005). Genome streamlining in a cosmopolitan oceanic bacterium. Science *309*, 1242–1245.

14. Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A.Z., Robbens, S., Partensky, F., Degroeve, S., Echeynié, S., Cooke, R., *et al*. (2006). Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. Proc. Natl. Acad. Sci. USA *103*, 11647–11652.

15. Cornillot, E., Hadj-Kaddour, K., Dassouli, A., Noel, B., Ranwez, V., Vacherie, B., Augagneur, Y., Brès, V., Duclos, A., Randazzo, S., *et al*. (2012). Sequencing of the smallest Apicomplexan genome from the human pathogen *Babesia microti*. Nucleic Acids Res. *40*, 9102–9114.

16. Haag, K.L., James, T.Y., Pombert, J.F., Larsson, R., Schaer, T.M., Refardt, D., and Ebert, D. (2014). Evolution of a morphological novelty occurred before genome compaction in a lineage of extreme parasites. Proc. Natl. Acad. Sci. USA *111*, 15480–15485.

17. Martin, W., and Koonin, E.V. (2006). Introns and the origin of nucleus–cytosol compartmentalization. Nature *440*, 41–45.

18. Martin, W., and Müller, M. (1998). The hydrogen hypothesis for the first eukaryote. Nature *392*, 37–41.

Gene Amplification: Trophoblast Giant Cells Use All the Tricks

James C. Cross
Department of Comparative Biology and Experimental Medicine, University of Calgary, Calgary, Alberta, T1S 1A2 Canada
Correspondence: jcross@ucalgary.ca
http://dx.doi.org/10.1016/j.cub.2015.12.030

Evolutionary gene duplication, developmental endoreduplication and selective gene amplification are alternative strategies for increasing gene copy number. When these processes occur together, things get really interesting, and new work shows that is the lifestyle of cells in the placenta.

Genomes are rapidly changing during evolution through mutations, as well as through gene-duplication events that create gene families. It has become clear that somatic cells also undergo genomic changes that are associated with new functions. Amplification of oncogenes is a common occurrence in cancer cells [1], and is presumably selected for by giving a growth advantage to cells. Endoreduplication is widely recognized during development of some cell types in plants, insects and mammals [2], and represents a more widespread amplification of genes, and results in polytene chromosomes. In mammals, trophoblast giant cells of the rodent placenta show extensive endoreduplication, doubling in DNA content for each round of DNA replication, and can reach a DNA content equivalent of over 1000 copies of the genome (Figure 1) [3,4], and it had been assumed for years that endoreduplication in trophoblast giant cells covered the entire genome. However, it was recently shown using new-generation genome analytical tools that, while the entire genome is polyploid, about 5% of the genome is relatively under-replicated [5]. In a report recently published in *Current Biology*, Roberta Hannibal and Julie Baker extend their analysis with even more sensitive methods and identify regions that are relatively over-replicated, representing less than 0.2% of the genome [6]. These regions include genes that are highly expressed in trophoblast cells but, amazingly, these regions encompass large multi-gene families that arose through gene duplication.

The mouse placenta contains several subtypes of trophoblast giant cells that differ in location within the placenta, cell lineage origin and the extent of polyploidy [7]. Hannibal and Baker focused their analysis on the parietal subtype of trophoblast giant cells, which are the first to form in the placenta, achieve the highest ploidies, and line the embryo implantation site and so are the easiest to isolate. Relative gene copy number across the genome was assessed at high resolution by using whole-genome sequencing, and then confirmed by digital droplet PCR, which is a probabilistic method for quantification, as it divides each PCR reaction into 20,000 droplets. Over-replicated parts of the genome occurred in five regions covering 4.7 million base pairs of the mouse genome. All five amplified regions contain gene families expressed in the placenta and thought to be important for normal pregnancy. One region is on mouse chromosome 6 and includes the NK/CLEC complex involved in interactions of natural killer cells in the