Convergence of Nutritional Symbioses in Obligate Blood Feeders

Olivier Duron, Yuval Gottlieb

Convergence of nutritional symbioses in obligate blood-feeders

Olivier Duron,[1,2*] and Yuval Gottlieb[3*]

[1] MIVEGEC (Maladies Infectieuses et Vecteurs : Ecologie, Génétique, Evolution et Contrôle), Centre National de la Recherche Scientifique (CNRS) - Institut pour la Recherche et le Développement (IRD) - Université de Montpellier (UM), Montpellier, France

[2] CREES (Centre de Recherche en Écologie et Évolution de la Santé), Montpellier, France

[3] Koret School of Veterinary Medicine, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, Israel

* Correspondances: olivier.duron@ird.fr (O. Duron); gottlieb.yuval@mail.huji.ac.il (Y. Gottlieb)

Key words: Hematophagy, Symbiosis, B vitamins, Biotin.

Abstract

Symbiosis with intracellular or gut bacteria is essential for the nutrition of animals with an obligate blood feeding habit. Divergent bacterial lineages have independently evolved functional interactions with obligate blood feeders, but all have converged to an analogous biochemical feature: the provisioning of B vitamins. Although symbionts and blood feeders coevolved interdependence for millions of years, we emphasize that their associations are not necessarily stable. Ancestral symbionts can be replaced by recently acquired bacteria with similar biochemical features. This dynamic emerged trough combination of phylogenetic and ecological constraints. Specifically, we highlight the lateral transfer of a streamlined biotin (B_7 vitamin) operon, and conjecture that its extensive spreading across bacterial lineages may drive the emergence of novel nutritional symbioses with blood feeders.

Symbiosis resolved key challenges in obligatory blood diet

Blood feeding (See Glossary) is one of the most specialized diets found in animals [1]. Blood is nutritionally unbalanced with high levels of protein, iron and salt, but few carbohydrates, lipids and vitamins. Blood feeder genomes evolved large repertories of genes related to vitamin and lipid shortage, haemoglobin digestion, iron managing or osmotic homeostasis to overcome these dietary challenges [1, 2]. Nevertheless, blood feeders cannot synthesize themselves essential cofactors and vitamins lacking in their diet [3, 4]. Facultative blood feeders, such as mosquitoes and fleas, usually need a blood meal to lay eggs but they also rely on other food sources over their life cycle and then avoid nutritional deficiencies. However, obligate blood feeders (OBF), as ticks, lice, leeches and vampire bats, cannot (**Box 1**). To overcome this constraint, OBF have converged to analogous functional **microbiomes** with **nutritional symbionts** able to synthesize several **B vitamins**.

An obligate blood feeding habit has independently emerged multiple times in animals including insects, arachnids, crustaceans, annelids and mammals, totalizing more than 7,400 species (**Box 1**). Accumulating studies demonstrate their ancient associations with B vitamin provisioning symbionts to condition the first appearance of OBF lineages that further radiated into current species [5-10]. The OBF microbiomes are functionally distinct from other animals: invertebrate OBF harbor typical low-complexity microbiomes (e.g. [10-14]), each dominated by one B vitamin provisioning symbiont (**Box 1**), while the only vertebrate OBF, vampire bats, harbor complex gut microbiomes with several potential B vitamin provisioning symbionts but that are distinct to microbiomes of insectivorous, carnivorous and frugivorous bats [15].

B vitamin provisioning symbionts as essential partners

Analogous B vitamins-based nutritional interactions appears to be strictly required for the survival and reproduction across the diverse OBF groups. We currently know little about vampire bats, as difficulties to maintain lab colonies make them fastidious experimental models. By contrast, experimental investigations on invertebrate OBF (including ticks, bed bugs, tsetse flies and kissing bugs) showed that, once deprived of their nutritional symbionts, they cease development, stop feeding, molting and reproduction (e.g. [13, 16-18]). They also exhibit physical abnormalities suggestive of a major vitamin deficiency, with dark and inflated bodies. Normal growth and development can be resumed only upon an artificial B vitamins supplementation or symbiont addition. Additional, albeit minor, contributions by nutritional symbionts exist, as exemplified in ticks by the production of the amino-acid L-proline [19].

B vitamin provisioning symbionts have evolved narrow associations with their hosts. They colonize only few organs of OBF, mostly **bacteriomes**, where they are hosted intracellularly in symbiotic cells termed **bacteriocytes**, but also gut caecae or **Malpighian tubules** in some species (**Figure 1, Box 2**). However, in triatomine bugs, symbionts live extracellularly in the lumen of the gut, although recent investigations also revealed complex microbiomes, including intracellular bacteria that could be additional B vitamin provisioning symbionts. In most cases, B vitamin provisioning symbionts are also heritable through successive generations via high fidelity maternal (usually transovarial, via oocyte infection) transmission (**Figure 1, Box 2**), and are further maintained during the life cycle of their hosts through **transstadial transmission**.

The emergence of B vitamin provisioning symbionts

B vitamin provisioning symbionts of OBF originated from at least two bacterial phyla (Proteobacteria and Actinobacteria) and within certain lineages of each phylum, notably the Enterobacteriaceae family (**Box 1 and Table 1**). We conjecture that their emergence in bacterial phyla depends on a combination of phylogenetic and ecological constraints through three mechanisms:

(i) The first mechanism implies that bacterial ancestors were already adapted to early mutualistic nutritional lifestyles, but with non-OBF organisms. Indeed, many Enterobacteriaceae are facultative symbionts, *ie*, non-required for host survival, but they are well known to influence animal nutrition and metabolism in diverse ways, and many are nutritional symbionts of non-OBF organisms. As such, some are extracellular symbionts inhabiting gut of vertebrates, others are intracellular symbionts of diverse arthropods, and most can produce B vitamins [20]. The combination of their broad distribution in animals and their biosynthesis capacity makes them a breeding ground for evolving nutritional symbiosis with OBF: the emergence of *Wigglesworthia* symbiont in tsetse flies, *Riesia* in lice and *Providencia* in leeches appears to occur through maintenance of ancestral B vitamin genes in these Enterobacteriaceae groups [6, 21, 22].

(ii) The second mechanism implies that bacterial ancestors were not adapted to a mutualistic nutritional lifestyle, but to a parasitic lifestyle with non-OBF organisms. This mechanism relies on maintenance of ancestral bacterial genes that encode for **pre-adaptations** to nutritional symbiosis, but that were primarily dedicated to another function, as best exemplified with the γ-proteobacterium

Francisella associated with ticks. It has emerged from a clade of virulent intracellular pathogens of vertebrates that includes the agent of tularemia *F. tularensis* [17, 23]. *Francisella* pathogens have evolved specific mechanisms to penetrate into phagocytes of mammals, and the self-production of biotin (B_7 vitamin) is here a key factor that enables pathogen replication and ultimate escape from the phagosomes [24, 25]. In the *Francisella* genus, the biotin biosynthesis pathway has evolved in the context of pathogenesis before being coopted for nutritional symbiosis in ticks.

(iii) The third mechanism depends on **lateral gene transfers**. Several B vitamin provisioning symbioses have independently evolved following biotin gene uptakes in the Rickettsiales order (Alpha-proteobacteria): *Wolbachia w*Cle in bed bugs, *Midichloria* in the castor bean tick *Ixodes ricinus* and *Rickettsia buchneri* in the black legged tick *I. scapularis* (**Figure 2**). All Rickettsiales are intracellular: some are pathogens, such as the agent of epidemic typhus *Rickettsia prowazekii*, while others are **reproductive parasites** of arthropods, such as *Wolbachia*, but only few harbour B vitamin genes [26-28]. Phylogenomic reconstructions revealed that three independent acquisitions of a streamlined biotin operon are at the origin of the Rickettsiales nutritional symbioses currently found in bed bugs and ticks [26-28]. While the acquisition of special 'symbiosis' genes is usually rare for nutritional symbionts [4, 29], the Rickettsiales nutritional symbioses show that foreign gene uptakes are key drivers of interactions with OBF. However, a different mechanism operates for folate (B_9): the folate biosynthesis pathway was early present in the Rickettsiales ancestor but secondarily lost in most sub-lineages

[30]. Remarkably, all the folate biosynthesis genes have been consistently maintained in Rickettsiales symbionts of bed bugs and ticks [26-28].

Convergence of B vitamin provisioning symbioses

Bacteria adapted to symbiotic (and specifically intracellular) lifestyles underwent massive genome reduction (**Box 3**). The gene set of B vitamin provisioning symbionts is largely a subset of the gene repertoires of their relatives: non-necessary genes have been pseudogenized or are missing completely, but B vitamin synthesis pathways have been conserved [6, 13, 17, 21, 22, 31-33]. However, depending on OBF symbiotic systems, certain B vitamin synthesis pathways have been maintained intact while others have been degraded or lost (**Table 1**). Notably, the symbiont genomes consistently harbour biosynthesis pathways of biotin and, at lesser extent, folate and riboflavin (B_2): these three pathways form a set of core genes fitting with the nutritional need of OBF. Each of the B vitamins are required for key enzymatic reactions in animals: biotin is a coenzyme for carboxylase enzymes, needed for fatty acids synthesis, branched-chain amino acid catabolism, and gluconeogenesis; folate is a precursor essential for the synthesis of DNA, the modification of DNA and RNA, and is also an important cofactor for cellular metabolism; riboflavin is a precursor of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) coenzymes, which are needed for a variety of flavoprotein enzyme reactions, including activation of other vitamins.

The presence of other B vitamin genes is more variable, and some pathways are missing one or more genes, or are entirely absent (**Table 1**). This pattern may actually depend on the symbiont lifestyle and the specific nutritional need of certain OBF. Indeed, the symbiont *Rhodococcus rhodnii* of kissing bugs has complete gene sets for the eight B vitamins (**Table**

1), suggesting that all are potentially needed either for kissing bug life cycle, or for *R. rhodnii* growth, or both. As an extracellular gut symbiont, *R. rhodnii* is exposed to fluctuating environments, in and out of host (it is transmitted by feces; **Figure 1**). This lifestyle is reflected in its large genome (4.3Mb) which exhibits important gene clusters dedicated to antimicrobial molecules and metabolic plasticity, and the eight B vitamin pathways may contribute to this plasticity [34]. By comparison, most B vitamin provisioning symbionts of OBF are intracellular; they live in a more stable, predictable and protected environment. As such, their metabolic needs are more limited and they have small genomes, lacking genes in almost all functional categories [29, 35, 36]. Hence, most intracellular symbionts have maintained intact the pathways for only few B vitamins other than biotin, ribflavin and folate (**Table 1**). However, some intracellular symbionts, as *Wigglesworthia* symbiont in tsetse flies and *Riesia* in lice, can produce most B vitamins (**Table 1**), suggesting that their hosts need this provisioning to their own growth and reproduction.

The fragile stability of B vitamin provisioning symbioses

B vitamin provisioning symbionts and OBF evolve a narrow interdependence, and each cannot survive without the other. Thanks to this **coevolution**, OBF nutritional symbioses have been traditionally envisioned as stable associations lasting for millions of years and resulting in **co-cladogenesis**, as shown in tsetse flies, and certain clades or genera of bat flies, louse flies, ticks and lice [5, 7-9, 13]. However, recent observations reveal that OBF nutritional symbioses are much more dynamic. Notably, nutritional symbioses can break down: recently acquired symbionts can replace ancestral B vitamin provisioning symbionts and provide similar benefits to the host. Such a pattern was observed in louse flies with the recent

acquisition of a *Sodalis* symbiont [38]. In sucking lice, there are up to six independent lineages of B vitamin provisioning symbionts [6, 7, 39-43], which suggests recent origin of each lineage and therefore several replacement events. In ticks, there are at least four independent lineages of B vitamin provisioning symbionts (*Coxiella, Francisella,Rickettsia* and *Midichloria*) [9, 16-17, 19, 23, 26, 28, 31, 33]. Notably, *Coxiella* symbioses are ancestral in some tick genera but recent replacements by *Francisella* in some species appear across the tick phylogeny [5, 9]. Genome sequencing otherwise confirmed that *Coxiella* and *Francisella* have roughly similar B vitamin biosynthesis capabilities: the recently acquired *Francisella* provides the same B vitamin benefit to ticks as the ancestral *Coxiella* (**Table 1**) [17, 19, 23, 31, 33].

Why ancestral and co-evolved nutritional symbionts are replaced in OBF remains unresolved, but several mechanisms can be proposed. Indeed, ancient symbionts suffer Muller's ratchet, with fixation of deleterious mutations through genetic drift, and they may just have over degraded genomes (**Box 3**). In this context, the comparison of OBF with sap-feeding insects is instructive: they host symbionts compensating for nutritional deficiencies of the sap diet. In sap-feeders, the most severely reduced of symbiont genomes are missing genes usually considered to be essential and harbor the tiniest known bacterial genomes. However, these nutritional symbioses do not collapse thanks to diverse mechanisms recently observed, including evolution of novel traits by hosts to compensate for symbiont gene losses, acquisition of another symbiont to supplement (or replace) functions that are lost in the older symbiont, or DNA uptake from environmental microbes to replace lost symbiont genes. These mechanisms are all potentially applicable to the symbioses with OBF. In ticks, such a pattern was observed for *Coxiella* with loss of essential genes for their replication, offering the

opportunity to another member of the microbiome to out-competing them [31, 33]. In addition, recently acquired symbionts may have higher biosynthetic capability than ancestral symbionts, and then supply additional benefits to OBF. In ticks, some ancestral *Coxiella* have genomes of only 0.66 Mb [33] while the recently acquired *Francisella* have bigger genomes (>1.5 Mb) that may have higher biosynthetic capability [17, 23]. This degeneration–replacement model has been proposed for other nutritional symbionts of arthropods [44-46], but are difficult to observe since replacements are expected to be transient [29]. However, the recent observation of a few tick species with co-infections by ancestral *Coxiella* and recently acquired *Francisella* may correspond to this transient state before extinction of the ancestral symbiont [5].

Invasion of a streamlined biotin operon

Accumulating genomic sequences confirmed that lateral transfer of a compact, streamlined, biotin operon is rampant in OBF nutritional symbioses: related biotin operons (i.e., that diverged recently from the same operon ancestor) were detected in diverse B vitamin provisioning symbionts of OBF (**Figure 2,** [26, 27, 43]). Related operons were also found in other intracellular bacteria, mostly in symbionts of non-OBF arthropods, such as the **reproductive parasite** *Cardinium* [47]. The incongruence between bacterial and operon phylogenetic trees underlines that these streamlined biotin operons experience recent (and likely ongoing) transfers between distantly related bacterial lineages (**Figure 2**). Its invasive nature may have contributed to major evolutionary innovations through the emergence of novel OBF nutritional symbioses. What is yet to be established are the mechanisms underlying the invasive nature of this streamlined biotin operon above other biotin operons.

These mechanisms may operate on different levels: the primary acquisition of the operon depends on the opportunity to DNA uptake from other bacteria, while its success to spread in the symbiont population rather depends on selection acting on the benefit it provides.

Lateral gene transfer is usually thought to be rare in intracellular symbionts as they reside in confined and isolated environments. However, according to the 'intracellular arena' hypothesis [48, 49], coinfections of different symbionts within the same host cell, and the propensity of some symbionts to switch between arthropod hosts, have created freely recombining intracellular bacterial communities [48-51]. The detection of the streamlined biotin operon in *Wolbachia*, *Rickettsia* and *Cardinium* [26, 27, 47] that are three of the most common intracellular symbionts of arthropods [52, 53], corroborates the 'intracellular arena' hypothesis and its role in emergence in novel OBF nutritional symbioses.

Once acquired, the further maintenance and spread of the operon in the symbiont population may be indicative of positive selection acting on it. A possibility is that the streamlined biotin operon is more efficient in producing biotin than others. In other operon systems, there are selective pressures for efficient specific gene orders and reduced intergenic regions to optimize the expression and functionality of operon in general [54, 55]. Others mechanisms may include the streamlined nature of this operon itself: its compact gene structure may favor its transfer in a single genetic block to other bacteria. Alternatively, the streamlined biotin operon may have genomic features favoring its transposition, but such a mechanism has not been detected to date.

Concluding remarks

That feeding specialization to strict blood diet is driven by nutritional symbioses is now beyond doubt. The capacity to synthesize B vitamins is widespread in bacteria but the nutritional symbionts of OBF have all converged to analogous, critical, interactions with their respective hosts. This convergence consists in severe degeneration of bacterial genomes accompanied by preservation of some B vitamin biosynthesis pathways, or in some cases, by a secondary acquisition of the streamlined biotin operon. The OBF nutritional symbioses can be ancient and highly co-evolved. However, we have now to consider that they are also influenced by a dynamic and complex web of interactions by which symbionts move between hosts and genes move between symbionts. We postulate that a better characterization of this web of interactions is now required if we are to understand the mechanisms driving the different aspects of nutritional symbiosis with OBF, such as convergence of B vitamin biosynthesis capacity, instability of association, extinction of ancestral symbiont, acquisition of novel symbionts or invasive spread of the streamlined biotin operon (see Outstanding Questions). Ecological opportunities, along with phylogenetic constraints and selective pressures acting on symbiotic systems, may altogether explain common and divergent evolutionary patterns in OBF nutritional symbiosis. In this context, comparative ecological and genomic approaches will be highly valuable in enhancing understanding of OBF nutritional specialty via symbioses.

Acknowledgments

This work has benefited from (1) an international ‘Joint Research Projects’ grant (EVOSYM) co-managed by the Ministry of Science, Technology and Space (Israel) and the Centre National de la Recherche Scientifique (CNRS, France), from (2) ‘Investissements d’Avenir’

grants (MITICKS, MICROBIOMES, GUYAVEC) managed by the Agence Nationale de la Recherche (ANR, France, Laboratoire d'Excellence CEBA, ref. ANR-10-LABX-25-01), and from (3) The Israel Science Foundation (ISF grant No. 1074/18).

References

1. Sterkel, M. et al. The Dose makes the poison: nutritional overload determines the life traits of blood-feeding arthropods. *Trends Parasitol.* 2017; 33: 633-644

2. Gulia-Nuss, M. et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nature Comm.* 2016; 7:10507

3. Gil, R. et al. Evolution of prokaryote-animal symbiosis from a genomics perspective. 2010: In *(Endo) symbiotic Methanogenic Archaea*, pp. 207-233, Springer.

4. Moran, N.A. et al. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet*. 2008; 42: 165-190

5. Binetruy, F. et al. Microbial community structure reveals instability of nutritional symbiosis during evolutionary radiation of *Amblyomma* ticks. *Mol. Ecol*. 2020; 29: 1016–1029

6. Boyd, B.M. et al. Genome sequence of *Candidatus* Riesia pediculischaeffi, endosymbiont of chimpanzee lice, and genomic comparison of recently acquired endosymbionts from human and chimpanzee lice. 2014; *G3 (Bethesda)* 4: 2189-2195

7. Boyd, B.M. et al. Primates, lice and bacteria: speciation and genome evolution in the symbionts of hominid lice. 2017; *Mol. Biol. Evol*. 34: 1743-1757

8. Chen, X. et al. Concordant evolution of a symbiont with its host insect species: molecular phylogeny of genus *Glossina* and its bacteriome-associated endosymbiont, *Wigglesworthia glossinidia*. *J. Mol. Evol*. 1999; 48: 49-58

9. Duron, O. et al. Evolutionary changes in symbiont community structure in ticks. *Mol. Ecol*. 2017; 26: 2905-2921

10. Hosokawa, T. et al. Reductive genome evolution, host–symbiont co-speciation and uterine transmission of endosymbiotic bacteria in bat flies. ISME J. 2012; 6: 577-587

11. Binetruy, F. et al. Surface sterilization methods impact measures of internal microbial diversity in ticks. *Parasit. Vectors.* 2019; 12: 268

12. Griffith, B.C. et al. Analysis of the gut-specific microbiome from field-captured tsetse flies, and its potential relevance to host trypanosome vector competence. *BMC Microbiol*. 2018; 18: 146

13. Hosokawa, T. et al. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proc. Nat. Acad. Sci*. 2010; 107: 769-774

14. Lalzar, I. et al. Composition and seasonal variation of *Rhipicephalus turanicus* and *Rhipicephalus sanguineus* bacterial communities. *Appl. Environ. Microbiol*. 2012; 78: 4110-4116

15. Mendoza, M.L.Z. et al. Hologenomic adaptations underlying the evolution of sanguivory in the common vampire bat. *Nat. Ecol. Evol*. 2018; 2: 659-668

16. Ben-Yosef, M. et al. *Coxiella*-like endosymbiont of *Rhipicephalus sanguineus* is required for physiological processes during ontogeny. *Frontiers Microbiol*. 2020; 11: 493

17. Duron, O. et al. Tick-bacteria mutualism depends on B vitamin synthesis pathways. *Curr. Biol*. 2018; 28: 1896-1902. e5

297 18. Michalkova, V. et al. Vitamin B6 generated by obligate symbionts is critical for
298 maintaining proline homeostasis and fecundity in tsetse flies. *Appl. Environ. Microbiol*. 2014;
299 80: 5844-5853
300 19. Tsementzi, D. et al. Comparison of closely related, uncultivated *Coxiella* tick
301 endosymbiont population genomes reveals clues about the mechanisms of symbiosis.
302 *Environ. Microbiol*. 2018; 20: 1751-1764
303 20. Husník, F. et al. Multiple origins of endosymbiosis within the Enterobacteriaceae (γ-
304 Proteobacteria): convergence of complex phylogenetic approaches. *BMC Biol.* 2011; 9: 87
305 21. Akman, L. et al. Genome sequence of the endocellular obligate symbiont of tsetse flies,
306 *Wigglesworthia glossinidia*. *Nat. Genet*. 2002; 32: 402-407
307 22. Manzano-Marín, A. et al. Solving a bloody mess: B-vitamin independent metabolic
308 convergence among gammaproteobacterial obligate endosymbionts from blood-feeding
309 arthropods and the leech *Haementeria officinalis*. *Genome Biol. Evol*. 2015; 7: 2871-2884
310 23. Gerhart, J.G. et al. Multiple acquisitions of pathogen-derived *Francisella* endosymbionts
311 in soft ticks. *Genome Biol. Evol*. 2018; 10: 607-615
312 24. Feng, Y. et al. A *Francisella* virulence factor catalyses an essential reaction of biotin
313 synthesis. Mol. Microbiol. 2014; 91: 300-314.
314 25. Napier, B.A. et al. Link between intraphagosomal biotin and rapid phagosomal escape in
315 *Francisella*. *Proc. Nat. Acad. Sci.* 2012; 109: 18084-18089
316 26. Gillespie, J.J. et al. A *Rickettsia* genome overrun by mobile genetic elements provides
317 insight into the acquisition of genes characteristic of an obligate intracellular lifestyle. *J. Bact*.
318 2012; 194: 376-394

27. Nikoh, N. et al. Evolutionary origin of insect *Wolbachia* nutritional mutualism. *Proc. Nat. Acad. Sci.* 2014; 111: 10257-10262

28. Sassera, D. et al. Phylogenomic evidence for the presence of a flagellum and cbb3 oxidase in the free-living mitochondrial ancestor. *Mol. Biol. Evol*. 2011; 28: 3285-3296

29. McCutcheon, J.P. et al. The life of an insect endosymbiont from the cradle to the grave. *Curr. Biol*. 2019; 29: R485-R495

30. Hunter, D.J. et al. The *Rickettsia* endosymbiont of *Ixodes pacificus* contains all the genes of de novo folate biosynthesis. *PloS One*. 2015; 10: 12

31. Gottlieb, Y. et al. Distinctive genome reduction rates revealed by genomic analyses of two *Coxiella*-like endosymbionts in ticks. *Genome Biol. Evol.* 2015; 7: 1779-1796

32. Perotti, M.A. et al. Host-symbiont interactions of the primary endosymbiont of human head and body lice. *FASEB J*. 2007; 21: 1058-1066

33. Smith, T.A. et al. A *Coxiella*-like endosymbiont is a potential vitamin source for the Lone Star tick. *Genome biol. Evol*. 2015; 7: 831-838

34. Pachebat, J.A. et al. Draft genome sequence of *Rhodococcus rhodnii* strain LMG5362, a symbiont of *Rhodnius prolixus* (Hemiptera, Reduviidae, Triatominae), the principle vector of *Trypanosoma cruzi*. *Genome announc.* 2013; 1: e00329-13

35. Moran, N.A. and Bennett, G.M. The tiniest tiny genomes. *Annu. Rev. Microbiol*. 2014; 68: 195-215

36. Bennett, G.M. and Moran, N.A. Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. *Proc. Nat. Acad. Sci.* 2015; 112: 10169-10176

37. Dadd, R. Nutrition: organisms. *Comprehensive insect physiology, biochemistry and pharmacology*. 1985; 4: 313-390

342 38. Šochová, E. et al. *Arsenophonus* and *Sodalis* replacements shape evolution of symbiosis in
343 louse flies. *PeerJ*. 2017; 5: e4099
344 39. Allen, J.M. et al. Effects of 16S rDNA sampling on estimates of the number of
345 endosymbiont lineages in sucking lice. *PeerJ*. 2016; 4: e2187
346 40. Fukatsu, T. et al. Intestinal endocellular symbiotic bacterium of the macaque louse
347 *Pedicinus obtusus*: distinct endosymbiont origins in anthropoid primate lice and the old world
348 monkey louse. *Appl. Environ. Microbiol*. 2009; 75: 3796-3799
349 41. Hypša, V. and Křížek, J. Molecular evidence for polyphyletic origin of the primary
350 symbionts of sucking lice (Phthiraptera, Anoplura). *Micro. Ecol*. 2007; 54: 242-251.
351 42. Říhová, J. et al. A new symbiotic lineage related to *Neisseria* and *Snodgrassella* arises
352 from the dynamic and diverse microbiomes in sucking lice. *bioRxiv*. 2019; 867275.
353 doi.org/10.1101/867275
354 43. Říhová, J. et al. *Legionella* becoming a mutualist: Adaptive processes shaping the genome
355 of symbiont in the louse *Polyplax serrata*. *Genome Biol. Evol.* 2017; 9: 2946-2957
356 44. Campbell, M.A. et al. Genome expansion via lineage splitting and genome reduction in
357 the cicada endosymbiont *Hodgkinia*. *Proc. Nat. Acad. Sci*. 2015; 112: 10192-10199
358 45. Łukasik, P. et al. Multiple origins of interdependent endosymbiotic complexes in a genus
359 of cicadas. *Proc. Nat. Acad. Sci*. 2018; 115: E226-E235
360 46. Matsuura, Y. et al. Recurrent symbiont recruitment from fungal parasites in cicadas. *Proc.*
361 *Nat. Acad. Sci*. 2018; 115: E5970-E5979
362 47. Penz, T. et al. Comparative genomics suggests an independent origin of cytoplasmic
363 incompatibility in *Cardinium hertigii*. 2012; *PLoS Genet*. 8: 10

48. Bordenstein, S.R. and Wernegreen, J.J. Bacteriophage flux in endosymbionts (*Wolbachia*): infection frequency, lateral transfer, and recombination rates. *Mol. Biol. Evol.* 2004; 21: 1981-1991

49. Kent, B.N. and Bordenstein, S.R. Phage WO of *Wolbachia*: lambda of the endosymbiont world. Trends Microbiol. 2010; 18: 173-181

50. Duron, O. Lateral transfers of insertion sequences between *Wolbachia, Cardinium* and *Rickettsia* bacterial endosymbionts. *Heredity* 2013; 111: 330-337

51. Duron, O. *Arsenophonus* insect symbionts are commonly infected with APSE, a bacteriophage involved in protective symbiosis. *FEMS Microbiol. Ecol*. 2014; 90: 184-194

52. Duron, O. et al. The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol*. 2008; 6: 27

53. Weinert, L.A. et al. The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc. R. Soc. B.* 2015; 282: 20150249

54. Lim, H.N. et al. Fundamental relationship between operon organization and gene expression. *Proc. Nat. Acad. Sci*. 2011; 108: 10626-10631

55. Wells, J.N. et al. Operon gene order is optimized for ordered protein complex assembly. *Cell Reports* 2016; 14: 679-685.

56. Lalzar, I. et al. Tissue tropism and vertical transmission of *Coxiella* in *Rhipicephalus sanguineus* and *Rhipicephalus turanicus* ticks. Environ. Microbiol. 2014; 16: 3657-3668

57. Nováková, E. et al. *Arsenophonus* and *Sodalis* symbionts in louse flies: an analogy to the Wigglesworthia and Sodalis system in tsetse flies. *Appl. Environ. Microbiol*. 2015; 81: 6189-6199.

386 58. Beard B. et al. Bacterial symbionts of the triatominae and their potential use in control of
387 chagas disease transmission. *Annu. Rev. of Entomol*. 2002; 47: 123-141
388 59. Graf, J. et al. Leeches and their microbiota: naturally simple symbiosis models. *Trends*
389 *Microbiol*. 2006; 14: 365-371
390 60. Lehane, M.J. et al. The biology of blood-sucking in insects. *Cambridge University Press*.
391 2005
392 61. Toft, C. and Andersson, S.G. Evolutionary microbial genomics: insights into bacterial
393 host adaptation. *Nat. Rev. Genet*. 2010; 11: 465-475

394

395 **Glossary**

396 **Bacteriocyte**: Specialized giant cell (also known as mycetocytes) of certain insects as tsetse
397 flies, bat flies, bed bugs but also aphids or weevils. It specifically contains endosymbionts
398 which provide essential amino acids and other molecules to their hosts, including B vitamins.
399 **Bacteriome**: Microbes containing cell clusters and surrounding cells which form altogether a
400 symbiotic organ. Depending on species, bacteriomes are found near the gut, or in the gut wall,
401 the gonads or nested among other organs within abdominal body cavity.
402 **Blood feeding**: The practice, either occasional or exclusive, of certain animals to feed on
403 blood of vertebrates. Also known as blood sucking, hematophagy or sanguinivory.
404 **B vitamins**: A group of chemically diverse water-soluble vitamins, commonly synthesized by
405 microorganisms, that are essential for cell metabolism and cannot be produced by animals.
406 The eight types of B vitamins are: thiamine (B_1), riboflavin (B_2), nicotinic acid, also known as

niacin (B_3), pantothenic acid (B_5), pyridoxine (B_6), biotin (B_7), folate (B_9) and cobalamin (B_{12}).

Co-cladogenesis: The parallel process of speciation between host and symbiont, such that phylogenetic trees of each partner are congruent.

Coevolution: The evolution of two or more species which reciprocally affect each other through the process of natural selection.

Lateral gene transfer (LGT): The movement of genetic material (including genes, operon, genomic islands, or more broad genomic regions) between unrelated organisms that may belong to different genera, families, orders, classes, phyla or even kingdoms, and that usually do not broadly exchange their genetic material. Also known as Horizontal gene transfer (HGT).

Malpighian tubules: Excretory and osmoregulatory organs of arthropods that lie in the abdominal body cavity and empty into the junction between midgut and hindgut.

Microbiome: The community of commensal, mutualistic and pathogenic microorganisms that reside in an environmental niche (such as an animal for instance).

Nutritional symbionts: microbes, usually bacteria and yeasts, that synthesize key nutrients lacking in the hosts' diet.

Pre-adaptations: The possibility of a certain characteristic trait to adopt a new biological function without evolutionary modification, or to have the same function in a different environmental context.

Reproductive parasite: Maternally inherited microorganisms that spread within arthropod populations by manipulating the host reproductive processes to enhance their own transmission. Manipulations involve biasing the sex ratio of infected females towards the

production of daughters through parthenogenesis, feminization of genetic males or male killing. Some reproductive parasites also induce reproductive incompatibility (cytoplasmic incompatibility) between males infected with a particular strain of bacteria and females not infected with this strain.

Transovarial transmission: Transmission via oocyte infection.

Transtadial transmission: Transmission during ontogeny from one life cycle stage to the next.

Box 1. Diversity of obligate blood feeders and their symbionts

Obligate blood feeding (OBF) habit has independently emerged multiple times in animals. There are more than 7,400 OBF species currently known worldwide [60]. Most of the OBF species are found in insects including:

● Sucking lice (Anoplura), with more than 5000 species. The Anoplura are all blood-feeding ectoparasites of mammals. A few species are parasites of humans, including the human body louse *Pediculus humanus humanus*, and the human pubic louse *Pthirus pubis*. Body lice are vectors for the transmission of the human diseases: epidemic typhus, trench fever, and relapsing fever.

● True bugs (Heteroptera), with at least 100 species of cimicids (bed bugs and relatives), 32 species of Polyctenidae bat bugs and more than 130 species of kissing (aka assassin, triatomine or vampire) bugs. Most species of cimicid are specialised on insectivorous bats or birds, while a few species, the common bed bug *Cimex lectularius* and its tropical relative *C. hemipterus*, feed on humans. Kissing bugs share shelter with nesting arboreal vertebrates, from which they suck blood. They are mainly found and widespread in the Americas where they are vectors of the Chagas disease parasite *Trypanosoma cruzi*.

● True flies (Diptera) with tsetse flies (>20 species), bat flies (>500 species) and louse flies (or keds, >150 species). They belong to the superfamily Hippoboscoidea and are obligate parasites of mammals and birds, often with a crab-like (louse fly) or a spider-like (bat fly) appearance. They reproduce through adenotrophic viviparity (see **Box 2**). Tsetse flies are vectors of trypanosomes, which cause human sleeping sickness and animal trypanosomiasis.

Other OBF species exist in arachnids:

459 ● Ticks (Acari: Ixodidea), with more than 900 species. Depending on species, they feed on
460 mammals, birds, but also reptiles and amphibians. Best-known tick-borne disease is the Lyme
461 disease that is nowadays a widespread infectious disease in the northern hemisphere.
462 ● Other Acari groups, with an undetermined number of species, such as the poultry red mite
463 *Dermanyssus gallinae* (vector of avian spirochaetosis), and the snake mite *Ophionyssus*
464 *natricis* (vector of ophidian paramyxovirus).
465 Only few other OBF species exist in non-arthropod groups:
466 ● Leeches (Hirudinea), with more than 700 species. The salivary glands of these annelids
467 produce an anticoagulant peptide, hirudin, used to treat some blood-clotting disorders.
468 ● Fish lice (crustacean subclass Branchiura), with approximately 120 known species. They
469 are the most widespread ectoparasites of freshwater fish in the world and can cause the severe
470 disease state argulosis. All life stages of both sexes are parasitic.
471 ● Vampire bats (Chiroptera: Desmodontinae), with only three species. All are native to the
472 Americas. Although rare, vampire bat bites can infect humans by rabies.

473 Phylogeny of representative B vitamin provisioning symbionts associated with these obligate
474 blood feeders is depicted in Figure I.

475

476 **Figure I. Evolutionary relationships of major examples for B vitamin provisioning**
477 **symbionts associated with obligate blood feeders.** The phylogeny is based on widely
478 supported findings from studies listed in the citations. B vitamin provisioning symbionts
479 associated with obligate blood feeders are shown in red. The bacterial phylogeny was drawn
480 from iTOL (https://itol.embl.de/).

481

Box 2. Localization and transmission of B vitamin provisioning symbionts

Reliable transmission mechanisms of B vitamin provisioning symbionts are necessary to stabilize nutritional associations between generations of OBF, and, more broadly, over evolutionary time. The efficiency of symbiont transmission can be enhanced by the evolution of co-adapted traits that ultimately lead to greater interdependence. Each lineage of OBF exhibits specific traits that condition what mechanisms of symbiont transmission is more favorable. In most OBF harboring **bacteriocytes**, such as bed bugs and lice, B vitamin provisioning symbionts are transmitted vertically through transfer from the maternal bacteriocyte to the ovaries of the female host, and thence to the eggs, a process known as **transovarial transmission** (see also **Figure 1**). In other obligate blood feeders harboring bacteriocytes, such as tsetse flies, bat flies and louse flies, the transmission of B vitamin provisioning symbionts arise latter during the insect development. Females of tsetse flies, bat flies and louse flies retain each egg within her uterus to have the offspring develop internally during the first three larval stages, a method called adenotrophic viviparity. During this time, the female feeds the developing offspring with a milky substance secreted by a modified gland in the uterus. The milky substance contains B vitamin provisioning symbionts which are then transmitted to the developing insect larvae. In triatomine bugs, the B vitamin provisioning symbionts show an exception among invertebrates as they are found in the gut, free of host cells and are transmitted to offspring via the feces either by egg shell contamination or by coprophagy, or even via cannibalism. The excretion of large quantities of water following the quick processing of the enormous volume of ingested blood, combine with the gregarious behavior of the triatomines, may have facilitated the evolution of this transmission mode. In vampire bats, we currently do not know how individuals acquire their

505 B vitamin provisioning symbionts, but it seems probable that transmission takes place at the
506 time of birth when a newborn is exposed to a mother's microbiota.

Box 3. Genomic decay in bacterial symbionts

Nutritional symbionts required for host survival harbor unusual genome modifications, including extreme reduction, rapid protein evolution, low GC-content and codon reassignments [4, 29, 36]. The process of genome reduction is initiated as a consequence of loss of selection on multiple gene functions when a free-living bacterium becomes host-associated [61]. The within-host environment, particularly inside a host cell, is relatively stable and nutrient rich, and the need for motility, regulation, secondary metabolite biosynthesis and defense is largely lost: over time the genes encoding such functions are pseudogenized and further definitely lost [4, 29]. For nutritional symbionts required for host survival, this process can continue even further, leading to tiny genomes where even genes considered essential are lost [4, 29, 36]. Symbiont genome decay however affects genes in all functional categories, even those involved in beneficial interactions with hosts [4, 29, 36]. Symbiont genomes continuously accumulate deleterious mutations and their degeneration may ultimately lead to maladaptation and then limit their beneficial contributions to their hosts. A main driving force for this process comes from strict clonality and small population size of nutritional symbionts during transmission, on which Muller's ratchet effect, leading to elevated rates of fixation of deleterious mutations irreversibly. Genetic drift favors the fixation of neutral or deleterious mutations that cause gene inactivation, gene loss, or inefficiency of gene products. Indeed, gene products of nutritional symbionts have lower efficiencies and reduced thermal stability than their homologs in free-living relatives. Once a bacterium has proceeded down the irreversible path into such obligate symbiosis, there is little opportunity to exit and the nutritional symbiosis can break down and collapse [29, 36].

Figure 1. Major examples for tissue tropism and mode of inheritance B vitamin provisioning symbionts associated with obligate blood feeders. Red dots: B vitamin provisioning symbionts in major hosting organ; Violet dots: B vitamin provisioning symbionts in gonads (for maternal transmission, excluding kissing bugs); Violet organs: gonads/milk glands (louse flies); Yellow organs: midgut/Malpighian tubules (ticks). EB- esophagus bacteriome; MGT- midgut; MGL-milk glands; MT-Malpighian tubules; OA-ovary ampule; OB- ovarian bacteriome; Ov- ovaries; SD- stomach disk; TB- testis bacteriome. Drawings are based on representative work of each taxon: Lice [32, 40]; ticks [17, 56]; Bedbugs [13]; Louse flies [57]; kissing bugs [58]; leeches [22, 59].

Figure 2. Evolutionary relationships, origin and structure of the streamlined biotin operon. The figure is based on widely supported findings from previous studies. Red, B vitamin provisioning symbionts associated with obligate blood feeders; Blue, intracellular symbionts of other arthropods; Black, intracellular pathogens of mammals. Filled arrows and white arrows indicate intact genes and pseudogenes, respectively.

Table 1. Biosynthetic pathways for B vitamins in 11 KEGG curated* genomes of symbionts associated with OBF.

B vitamin provisioning symbiont				Obligate blood feeder host	KEGG code	Vitamin B_1 Thiamine	Vitamin B_2 Riboflavin	Vitamin B_3 Nicotinic acid	Vitamin B_5 Pantothenic acid	Vitamin B_6 Pyridoxine	Vitamin B_7 Biotin	Vitamin B_9 Folate	Vitamin B_{12} Cobalamin
Phyla	Order	Family	Strain										
γ-proteobacteria	Legionellales	Coxiellaceae	*Coxiella* str. CRt	Tick (*Rhipicephalus turanicus*)	cey								
			Coxiella str. CeAS-UFV	Tick (*Amblyomma sculptum*)	cend								
			Coxiella str. CLEAA	Lone star tick (*Amblyomma americanum*)	cea								
γ-proteobacteria	Thiotrichales	Francisellaceae	*Francisella persica*	Soft tick (*Argas arboreus*)	fper								
γ-proteobacteria	Enterobacterales	Erwiniaceae	*Wigglesworthia glossinidia*	Tsetse fly (*Glossina brevipalpis*)	wbr								
			Wigglesworthia glossinidia	Tsetse fly (*Glossina morsitans*)	wgl								
γ-proteobacteria	Enterobacterales	Morganellaceae	*Arsenophonus lipopteni*	Deer keds (*Lipoptena fortisetosa*)	asy								**
γ-proteobacteria	Enterobacterales	Enterobacteriaceae	*Cand.* Riesia sp. GBBU	Gorilla louse (*Pthirus gorillae*)	rig								
			Cand. Riesia pediculicola str. USDA	Human body louse (*Pediculus humanus humanus*)	rip								
α-proteobacteria	Rickettsiales	Midichloriaceae	*Cand.* Midichloria mitochondrii	Castor bean tick (*Ixodes ricinus*)	mmn								
α-proteobacteria	Rickettsiales	Anaplasmataceae	*Wolbachia* str. *w*Cle	Bed bug (*Cimex lecturalis*)	wcl								

*KEGG pathways used for B_1, B_2, B_3, B_5, B_6, B_7, B_9 and B_{12} are 00730, 00740, 00760, 00770, 00750, 00780, 00790, 00860, respectively. Black squares, putatively functional pathways; Grey squares, incomplete pathways with pseudogenes or missing genes; white squares, pathways absent.

** Only flavin reductase (*fre*) present. An enzyme that converts Aquacob(III)alamin into vitamin B_{12}.